University of Naples “Federico II”

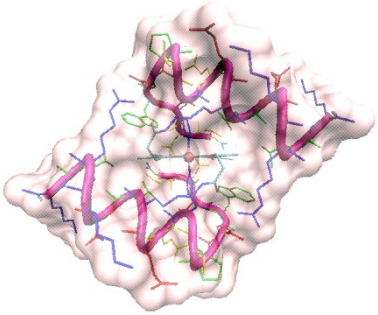
Department of Chemical Sciences

Ph.D. in Chemical Sciences

XXVI Cycle

Ph.D. student: **Claudia Vicari**

“Structural characterization of heme-protein models”



Advisors:

Prof. Vincenzo Pavone

Prof. Flavia Nastri

Examiner:

Prof. Gerardino D’Errico

Coordinator:

Prof. Luigi Paduano

Table of contents

Abstract

Metalloproteins play important roles in biology, as they account for nearly half of all proteins in nature, spanning a remarkable functional range, from fundamental chemical reactions, such as electron transfer, dioxygen transport and activation, hydrolysis, group transfer and rearrangements, to signaling processes governing gene regulation and expression.[1,2]

The same metal cofactor is able to serve different roles depending on the protein architecture, as the physical and chemical properties of a selected metal depends on the protein environment. In fact, a plethora of mixed interactions occur between the metal cofactor and the protein matrix, fine-tuning the intrinsic reactivity of the metal.[3]

A prominent illustration of how the metal center reactivity is modulated by the surrounding protein environment is provided by heme-proteins, that are a family of important macromolecules ubiquitous in biological systems characterized by a common prosthetic group (the "heme" group), which perform a wide range of biological functions, and participate in a wide range of essential biological processes: myoglobin and hemoglobin reversibly bind dioxygen, cytochromes support electron-transfer processes, cytochrome P450 activates dioxygen to form water and an oxygen atom, which is inserted into a substrate.[4]

A powerful approach to deeply understand the factors that specify the functions of natural heme-proteins is the engineering of peptide-based models that mimic protein active sites in simplified scaffolds. To date, several artificial heme-proteins have been engineered, providing provide further insights on -activity relationships.[5]

In this respect, the Artificial Metalloenzyme group, where this project has been carried out, developed a class of artificial heme-mimetics, named Mimochromes. These models are composed of two peptide chains covalently linked to the heme group via amide bonds between the porphyrin propionates and the lysine side-chains.[6]

Early compounds of this class were made up of two identical nona-peptide chains, each containing a histidine residue which acts as axial ligand to the heme central ion (iron or cobalt), in a low-spin bis-His-ligated state. The peptide chains were in an α-helical conformation in both the apo and metal containing species, and these molecules adopt a stable and well defined conformation, being able of supporting electron transfer processes.

More recently, many efforts have been made in order to design new compounds (Mimochrome VI and its analogues) that bind iron in five-coordinate complexes, with a vacant coordination site available for catalyzing oxidation reactions. This efforts lead to the development of new models that have peroxidase-like activity, being able of catalyzing organic substrate (such as ABTS) oxidation, by the use of hydrogen peroxide. In particular, one of this analogues, showed catalytic performance approaching that of HRP.[7]

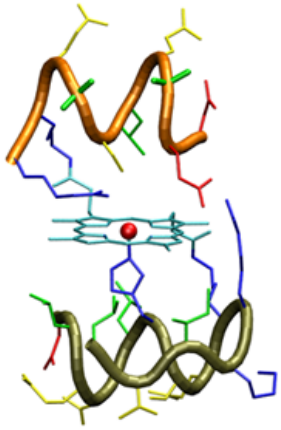
One main drawback of these new molecules relies in the lack of detailed information of their three-dimensional structures. Attempts to solve the structure on the fully diamagnetic Co(III) complex were unsuccessful. In fact, analysis of the Nuclear Magnetic Resonance (NMR) spectrum of Co(III)-Mimochrome VI revealed the presence of different species in solution.

This PhD project was aimed at overcoming the presence of multiple species in solution, developing heme-mimetics that couple high activity with a well defined structure. In this respect, two different strategies have been undertaken:

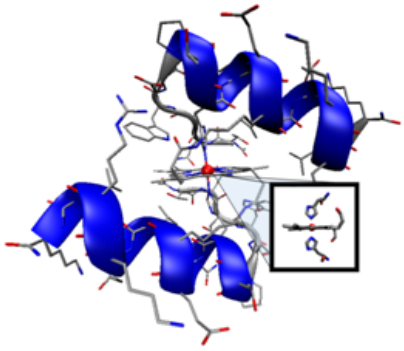
- incorporation of non-coded amino acid with α-helix stabilizing effect;

- elongation of the peptide chains, that allows the introduction of a large number of intra-chain and inter-chain interactions.

The first strategy led to the development of a new Mimochrome VI analogue, named Mimochrome VI-2U1L. The second strategy resulted in the design of the Mimochrome VII molecule.



Mimochrome VI-2U1L



Mimochrome VII

For both compounds, a detalied NMR analysis was performed, and information about the three- dimensional structure of the Co(III)- complexes was achieved.

For Co(III)-Mimochrome VI-2U1L, four different species were present in soultion. The four species were purified by RP-HPLC and separately analyzed by NMR. In particular, for two of these species (**peak a** and **peak b**), the solution structures of the tetradecapeptide-porphyrin moiety was solved. The tetradecapeptide adpots in solution the expected helical conformation in the Asp^{1} -Lys^{9}segment, while distortion is observed at the C-terminal region (intended to fold in extended conformation).

Regarding the decapeptide chain, severe overlaps make difficult the structural characterization. In order to solve the oserved ambiguity of assignment, further experiments are required. Indeed, the information obtained from the Co(III)-Mimochrome VI-2U1L molecule, will be very useful in order to analyze the structure-function relationship on Mimochorme VI and its analogues.

Regarding the Mimochrome VII molecule, two species were presnt in solution. The structure of the more abundant specie was determined and good agreement was observed between the designed model and the determined NMR structure, apart from the first 4 residues at the N-terminal region. The two peptide chains adopt right-handed helical conformation from Pro^{6} to Leu^{16}, while a less regular conformation is observed in the N-terminal region. As designed, the global structure is stabilized by the presence of interactions between hydrophobic side chains and deuteroporphyrin ring.

These information will be used to re-design new molecules characterized by asymmetrical metal binding site, in order to introduce new

functionalities. Preliminary, in this PhD thesis the design of a penta-coordinated model, Mimochrome VII asym, was carried out. This new model is characterized by a distal cavity that contains an arginine residue, mimicking Arg^{38} of Horseradish Peroxidase (HRP), and a non-coordinating His residue to mimic His^{42} of HRP. Its characterization is actually under course.

Finally, part of this work was devoted to the structural characterization of active sites in paramagnetic iron(III) mimochrome complexes. In particular, the orientation of the histdine axial ligands in the Fe(III) complexes of mimochrome IV was determined, by taking advantage of a methodology previously used for natural ferriheme-proteins.

Introduction

1. Chapter 1: design of artificial metallo-proteins

1.1. Metallo-proteins

Many life-sustaining process, like nitrogen fixation, photosynthesis and respiration, involve proteins containing metal-binding centers, that are generally referred as metallo-proteins. They account for nearly half of all proteins in nature, and play important roles in biology, spanning a remarkable functional range, from fundamental chemical reactions, such as electron transfer, dioxygen transport and activation, hydrolysis, group transfer and rearrangements, to signaling processes governing gene regulation and expression.[4,5,8,9] Despite their low abundance in terms of mass, numerous metal ion, including various 'trace' elements, are an absolute requirement for all life forms. Na^{+}, K^{+}, $Mg^{2+,}$ Ca^{2+}, $Zn^{2+,}$ Mn^{2+}, Ni^{2+}, $Cu^{+/2+}$, $Fe^{2+/3+}$, and $Co^{2+/3+}$ are frequently found in biological organisms.[8,10-19]Also, recent advances in mass spectrometry, X-ray scattering/absorption, and proteomics have demonstrated that even elements such as uranium, arsenic, and lead are involved in biological processes (**Fig. 1**). Interestingly, some metals can be toxic at elevated levels, but are essential in trace amounts.[20]

While a small amount of metals are present in the human body both as free cations (such as Na^{+} and K^{+}) or minerals (such as hydroxyapatite crystals), many others are incorporated into biomolecules as cofactors.[21] The physical and chemical properties of a selected metal depends on the protein environment, where a plethora of mixed interactions that occur between the metal cofactor and the protein matrix mutually affects the properties of each other, thus enhancing, diversifying, or tuning their individual functions.[22-25] Therefore, often the same metal cofactor is able to serve different roles depending on the protein architecture.

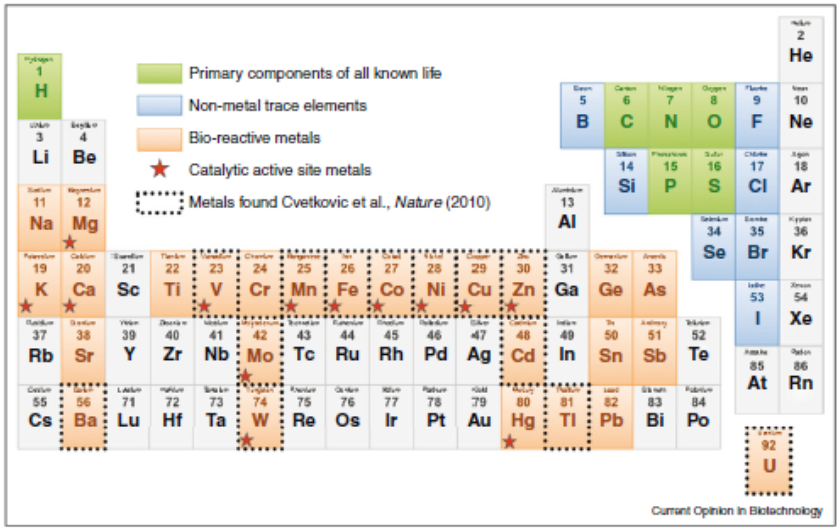


Fig. 1 Periodic table emphasizing biologically relevant elements and bio-metals. The highly abundant building blocks of nature are shown in green, metals described to be associated with metalloproteins are shown in orange and non-metal trace elements in blue. Catalytic metals present in active sites of enzymes are marked with a red star. Dotted boarders highlight metals incorporated into microbial biomass into a recent studies of Cvetkovic et al. Reprinted from *Current Opinion in Structural Biology*, Vol. 23, Yannone et al., Metals in biology: defining metalloproteomes, pages 89-95.,[20]Copyright (2012), with permission from Elsevier.

Most metal-binding sites include only a small number of coordinating amino acids, in a specific spatial arrangement. In most cases, the coordinating amino acid are histidine (His), cysteine (Cys), aspartic (Asp), and glutamic acid (Glu). Less commonly, tyrosine (Tyr), methionine (Met), serine (Ser), lysine (Lys), glutamine (Gln), asparagine (Asn), threonine (Thr), and glycine (Gly) are also coordinated to metal ion (**Fig. 2**). With the exception of glycine, which coordinates a metal ion through the backbone carbonyl oxygen, the other amino acids participate in coordination almost exclusively through their side chains containing sulfur, oxygen or nitrogen as donor atoms.[11,3] Metal ion preferences for a certain amino acid side chain follow the general rules of coordination chemistry and often can be

interpreted using the theory of hard and soft acids and bases (HSAB theory).[26] Furthermore, metal ion can also be coordinated by exogenous non-amino acid cofactors, such as organic macrocyclic ligands, like porphyrin, corrin, and chlorin groups or inorganic entities (e.g., oxide, hydroxide, water).[27,28]

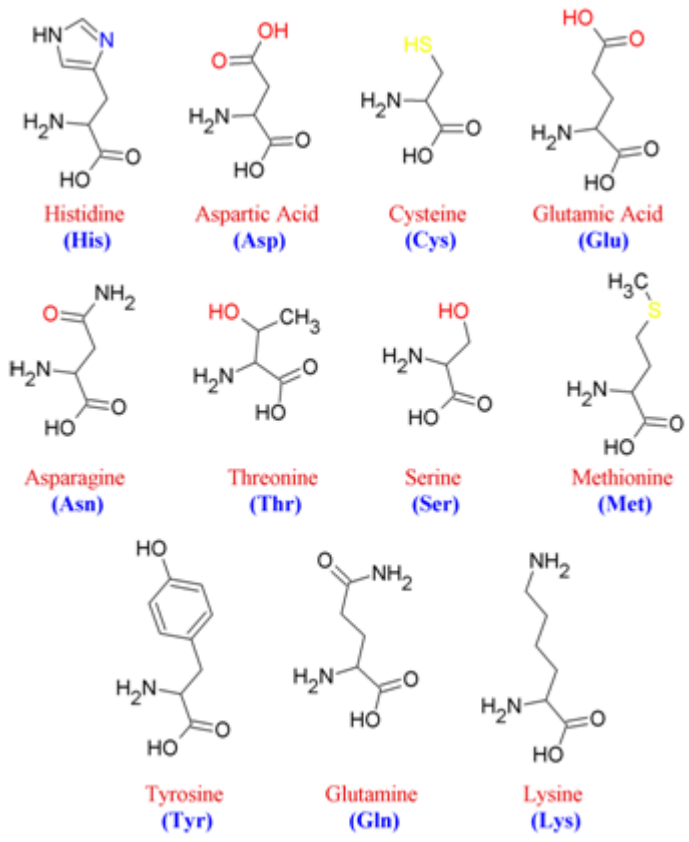


Fig. 2 Amino acids containing donor atom(s) in their side chain.

The ligand binding mode is a important factor in modulation of the protein affinity/selectivity for a given metal, and plays an important role in tuning the metal reactivity. Therefore, it is a key aspect in determining metallo-protein functions. However, metallo-protein functions are also subtly related to the specific arrangement of non-coordinating amino acid in

the proximity of the active site. Hydrogen bonds and ion pairs formation, charge distribution and specific ligand affinity strictly modulate the properties of the metal-binding site. Consequently, understanding the minimal requirements for function in metallo-protein is a complex task. Several approaches have been used to shed light on structure-function relations in metalloproteins:

- Bioinformatics, that aims at identifying metallo-proteins in diverse genomes from primary sequence data. However, this approach is generally limited to the analysis of closely homologous of well-characterized proteins. While various metal-binding motifs have been identified, and some are useful in classify probable metallo-proteins from primary sequence, in most cases computational approaches fail to predict specific protein-bound metals.[20]

- Structural biology, that aims to determine the three-dimensional structures of proteins. Structure databases are the main sources of information on the characteristics of metalloproteins at an atomic level. The Protein Data Bank (PDB), a worldwide archive of atomic coordinates and other information describing proteins and other important biological macromolecules, include about 30000 metallo-proteins structure, obtained by X-ray crystallography, Nuclear Magnetic Resonance (NMR) spectroscopy, and cryo-electron microscopy (**Fig. 3**).[29,30]

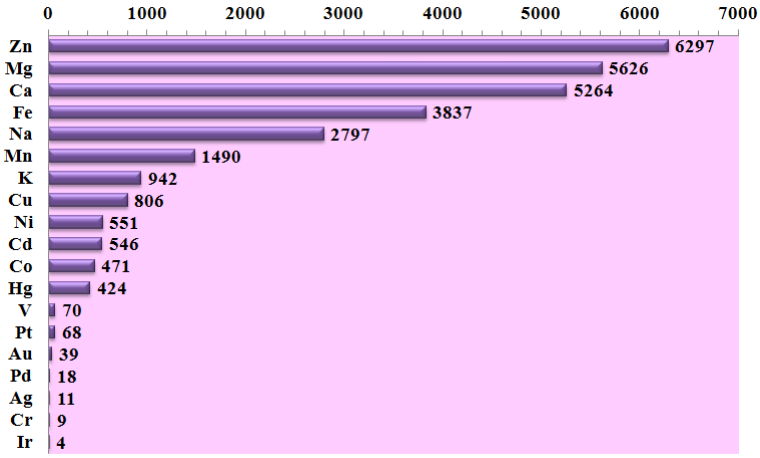


Fig. 3 Metals in the Protein Data Bank, January 2011.

- Mass spectrometry of metallo-proteins that is a powerful tool for analyzing many structural and behavioral aspects of metallo-proteins in great detail. Several MS techniques have been developed for determining the locations of metal-binding centers, metal oxidation states and reaction intermediates of metallo-proteins.[31-33] As example, inductively coupled plasma mass spectrometry (ICP-MS) allows rapid and sensitive quantification of metals from complex biological samples, ESI-MS enables the determination of protein–metal ion binding stoichiometry in solution, and metallo-protein higher order structure in the case of multi-subunit proteins.[20,21]

These and other approaches led to a deep understanding of the roles of metals in biological processes.[19] Traditionally, metallo-proteins are ranked with respect to their function in three classes (**Fig. 4**):

- transport or storage proteins;
- signal transduction proteins;
- metallo-enzymes.

The first class include metal storage proteins, electron carriers, and oxygen binding proteins. The second class involves metal sensing and switching proteins. The third class involves metallo-proteins able to catalyze a wide range of chemical reactions. Analysis of enzyme mechanisms, restricted to enzymes with known structure, has shown that metallo-enzymes represent about 40% of all enzymes.[34,35]

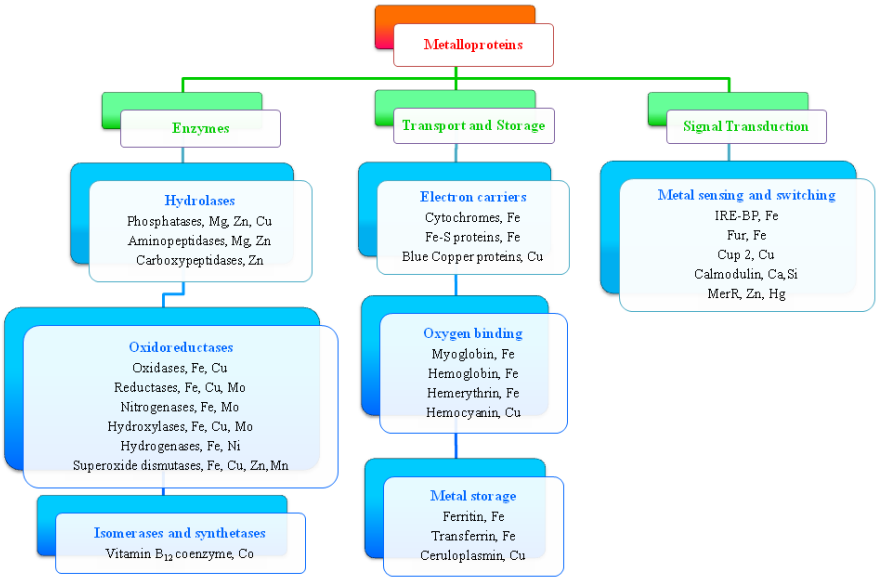


Fig. 4 Classification of metallo-proteins according to their functions.[35]

In recent years, the growing number of metallo-proteins of known three-dimensional structure have prompted numerous studies that attempt to classify metallo-proteins from a structural point of view according to metal-binding site properties, such as the presence of a certain metal, the coordination number and the amino-acid residues involved in metal coordination. All these studies provide new insights on protein folding, structure, and function of metallo-proteins, and are a valuable tool for the identification of metal-binding sites in proteins of unknown structure. Also,

they laid the groundwork for the design of artificial proteins that reproduce metal coordination environments of native proteins.

1.2. Heme-proteins

Heme-proteins are a class of biologically important macromolecules ubiquitous in biological systems and characterized by a common prosthetic group, the "heme" group: a tetrapyrrole macrocycle, called porphyrin, that bind an iron ion. Despite sharing the same cofactor, heme proteins perform a large array of biological functions, providing a clear illustration of how the metal center reactivity is modulated by the surrounding protein environment. In fact, depending on the heme environment, these proteins catalyze numerous chemical reactions, and participate in a wide range of essential biological processes: myoglobin and hemoglobin reversibly bind dioxygen, cytochromes support electron-transfer processes, cytochrome P450 activates dioxygen to form water and an oxygen atom, which is inserted into a substrate. Peroxidases use hydrogen peroxide to oxidize several organic substrates (**Fig. 5**).[4]

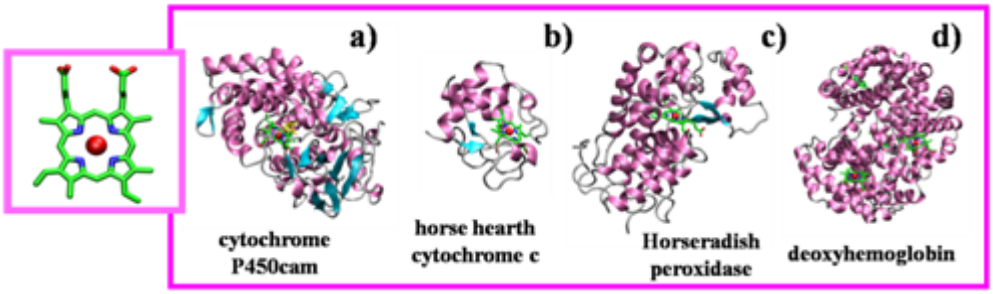


Fig. 5 Crystal structure of: **a)** cyochrome -P450cam (PDB code 5CP4);[36] **b)** horse hearth cytochrome c (PDB code 1AKK);[37] **c)** horseradish peroxidase (HRP) (PDB code 1ATJ);[38] **d)** deoxyhemoglobin (PDB code 1HGA).[39]

Due to this high functional versatility, heme-proteins have been the focus of numerous studies, aimed at describing the structural and functional

properties of heme binding sites. Myoglobin was the first protein whose structure was determined, using X-ray crystallography, by Sir John Cowdery Kendrew, an English biochemist and crystallographer, in 1958.[40] Two years later, Max Perutz published the structure of hemoglobin.[41] These discoveries earned to the two scientists the Nobel Prize in Chemistry 1962, that was awarded jointly to Max Ferdinand Perutz and John Cowdery Kendrew *"for their studies of the structures of globular proteins"*.[42] Since then, high-resolution X-ray crystallographic data, as well NMR three-dimensional structures, become available for an increasing number of heme-proteins, together with a large number of functional characterizations, and today the protein data bank contains over 2300 structures of proteins containing one or more heme groups.[43]

1.3. The heme prosthetic group

The basic structure of natural heme proteins is a tetrapyrrole macrocycle called porphyrin, with a skeleton as shown in **Fig. 6** that consists of four pyrrole-like five-membered units, which are linked by four methine bridges (meso positions). It has 11 double bonds, which means 22 π electrons, and 18 of these electron are involved in a delocalization pathway, forming an aromatic system that obeys Huckel's rule of aromaticity ($4n + 2$ π electrons, $n = 4$) and is shown by X ray crystallography to be planar. The porphyrin free base has two inner pyrrole protons that can be replaced by a metal ion (Fe, Co, Ni, Cu, Zn), forming a metallo-porphyrin complex, with the four pyrrole nitrogen atoms acting as metal ligands.

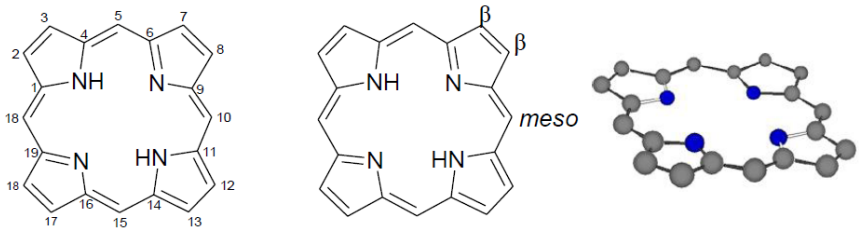


Fig. 6 The tetrapyrrole macrocycle.

The pyrrole nitrogen forms a relatively rigid hole for a cation, close to but smaller than the diameter of high-spin Fe^{2+}or Fe^{3+}. The porphyrin ligand is a good four-centered σ donor, and it is also an effective π acceptor. The energy level diagram of the iron porphyrin also indicate that the high-spin iron in the ring, whether in the Fe^{2+} or Fe^{3+} state, is quite close in energy to the low-spin state, and indeed to intermediate spin state. This spin state equilibrium balance is one of the delicate controls, which the porphyrin ring, together with the protein matrix, exhibits over the metal ion.[3]

The porphyrin macrocycle exists in various forms, differing in the pyrrolic substituents. The most common metalloporphyrin in biologically system is the heme b, an iron(II) protoporphyrin IX that noncovalently binds to proteins, found in b-type cytochromes, globins, cytochromes P-450 and hemesensor proteins. Another very common form is the heme c, a modified heme b in which the porphyrin macrocycle vinyl groups forms thioether bonds with two cysteine residues, forming a classical Cys-Xaa-Xaa-Cys-His (CXXCH) motif, found in cytochromes c and cytochrome f. Other heme types, obtained by modifications of b-type heme that can be found in heme-proteins are a- and d-type hemes contained, for example, in the terminal oxidases (heme a) and catalase (heme d).[5,44] The structures of these commonly occurring hemes are illustrated in **Fig. 7**.

In addition to the pyrrole nitrogen atoms, the metal ion is usually capable of binding one or two supplementary ligands, called axial ligands, that severely influence the spectroscopic, electrochemical and functional properties.[45] Several amino acids are able to serve as axial ligands to heme iron.

In some heme-proteins, the protein provides only one axial ligand, whereas the opposite side is open for the binding of small molecules (as example O_2 in globins) or organic substrates (cytochrome P450, peroxidases, catalases).

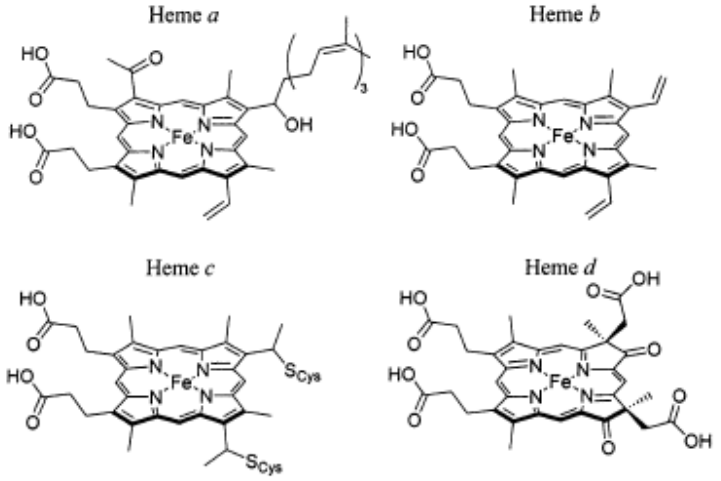


Fig. 7 Chemical structures of commonly occurring natural hemes a, b, c, and d. Reprinted with permission from Reedy and Gibney, Chem. Rev., 2004, 104, 617-649. Copyright (2004) American Chemical Society.

1.4. Heme binding to proteins

In heme-proteins, the heme group is linked to the protein matrix by a plethora of different interactions:

- coordination bond between the heme iron and one or two amino acids that act as axial ligands to the metal ion;
- hydrophobic interactions between non-polar amino acid and the porphyrin macrocycle;
- polar interactions between the propionic groups and polar side-chains;
- covalent bonds (in c-type heme), between the vinyl groups of the porphyrin macrocycle and cysteine residue side-chains.

Furthermore, an intricate set of different interactions also contribute to imprinting the heme properties. Hydrophobic, ion pairing, and hydrogen bonding interactions control several features of the heme environment, such as the local dielectric constant, the heme exposure, and the orientation of the ligands. Variations in the mode of heme binding, in the number, type, and donor properties of the heme axial ligands, and in non-binding interactions between the heme and the protein matrix, result in different heme environments in proteins with different functions in biology.

The characteristics of the primary coordination sphere, like metal coordination geometry, number, type, and donor properties of the axial ligands, play important roles in modulating the properties of heme proteins, such as redox potentials, electronic structure, spin states, electron-transfer rates, and catalytic properties. Several amino acids are known to serve as axial or proximal ligands to heme proteins.[43]

Table 1. Analysis of axial ligands for heme group on a data set of 34 nonhomologous heme proteins. Data from ref. 43.	
Axial ligands for heme group	
His	**61**
Met	**8**
Cys	**6**
Tyr	**3**
Pro	**1**
N-terminal amino	**1**
Vacant/water	**18**
Changes on oxidation/reduction	**2**

The primary coordination sphere of the iron in heme proteins of known structure is dominated by histidine, that ligates heme iron via the Nε, also if there is only a single example of N_δ-bound histidine (His^{102} of Nitrosomonas europea cyt c554).[46] In particular, the most frequently observed coordination number and ligation motif is five-coordinate mono-histidine, found as example in globin and peroxidases. In these proteins, the matrix provides only one axial ligand, histidine, whereas the opposite side is open for binding of exogenous molecules. The second most common coordination motif is six-coordinate bis-histidine, that is found in electron carriers proteins, and appear one-third as often as mono-histidine. Methionine and cysteine are also common axial ligands that coordinate the heme. In particular, a common ligation motif for six-coordinate heme proteins is histidine with methionine, as exemplified by cyt c.[47] Bis-methionine coordination is also observed (bacterioferritin).[48] Conversely, cysteine thiolate ligation is observed in five-coordinate heme proteins, such as cytochromes P-450 and nitric oxide synthase.[49,50] Furthermore, other amino-acid ligands are rarely found in six-coordinated heme-proteins in combination with histidine, like the side-chain amine of lysine (cyt c nitrite reductase)[51] the asparagine amide donor (ferric

SHP, Sphaeroides heme protein)[52], the charged donors tyrosine phenoxide (HasA)[53] and glutamate carboxylate (hemoglobin Milwaukee).[54] Finally, the neutral N-terminal amine donor has also been found coordinated to heme iron in few structures (Tyr^1 of cyt f and Pro^2 of ferric CooA).[55,56]

As well as the number and type of axial ligands, the coordination geometry play an important role on imprinting heme properties, the orientation of the axial ligands, with respect to the porphyrin plane, has been suggested to be one of the key factor in determining the heme features. This hypothesis originated from the evidence that, even within a common set of axial ligands and heme substituents, the physicochemical properties of heme proteins vary widely. In particular, it has been demonstrated that different histidine conformations can shift the redox potential of heme proteins,[46] and that the spectroscopic, electronic, and magnetic properties of ferriheme model complexes can depend on the orientation of axially coordinated planar ligands.[57]

While the axial ligands to heme iron establish the basic coordination chemistry, the interactions of amino acids beyond the primary coordination sphere (hydrophobic, ion pairing, and hydrogen bonding interaction) controls several features of the heme environment, such as the local dielectric constant, the heme exposure, and the ligand orientation, allowing it to perform a variety of biochemical functions. Further, the protein directs long-range interactions and regulates the accessibility of solvent and substrates into the active site, thus allowing to selectively discriminate between different ligands and/or substrates. These factors all contribute to the functional versatility of the heme. In the following sections, the focus is pointed on two heme-protein classes: peroxidase, belonging to the mono-histidine type and cytochromes belonging to the bis-histidine. A detailed

discussion of the main factors that characterize the active site of these proteins is given, in order to develop new artificial heme binding proteins, that reproduce their heme environment.

1.5. Peroxidases

Peroxidases are ubiquitous enzymes that catalyze the oxidation of a wide range of organic and inorganic substrates, such as phenols, aromatic amine, thioanisoles and iodide, using hydrogen peroxide.

They are crucial for the regulation of reactive oxygen species levels and for the promotion of several substrates oxidation, also play major roles during defence and development processes, and are among the first enzymes to have been discovered, with references dating back to the 19th century, describing peroxidase activity in biological system. In 1855, Schönbein observed the oxidation of some organic compounds, such as guaiacol, by hydrogen peroxide (H_2O_2), catalyze by "substances" present in plants and animals. In 1898 Linosseir used for the first time the name "peroxidase", describing an enzyme isolated from pus.[58] Since then, peroxidases are one of the most extensively studied groups of enzymes, and several heme containing peroxidases have been isolated, sequenced, and structurally characterized.

To date, almost 10000 heme peroxidase have been annotated and classified in the peroxidase database (*PeroxiBase,* http://peroxibase.isb-sib.ch).[59] Of these, more than 8000 belong to the "plant peroxidase" superfamily, which contains monomeric proteins from both prokaryotic and eukaryotic origin with a non-covalently bound heme. The plant peroxidases are further classified in three classes: class I that contains intracellular prokaryotic peroxidases, class II, that contains extracellular

fungal peroxidases, and class III that contains secretory plant peroxidases. The remaining heme peroxidase, are usually formed by association of different peptide chain. In particular, animal peroxidase are usually dimeric (with a covalently attached heme group), while catalase are tetrameric enzymes (that catalyze the dismutation of H_2O_2).

Even if the structures of several peroxidases have been determined, the majority of the information on the structure–activity relationship have been obtained from the analysis of yeast cytochrome c peroxidase (CCP, belonging to class I of plant peroxidases) and of Horseradish (Armoracia rusticana) peroxidase (HRP, belonging to class III of plant peroxidases) (**Fig. 8**).[58, 60,61]

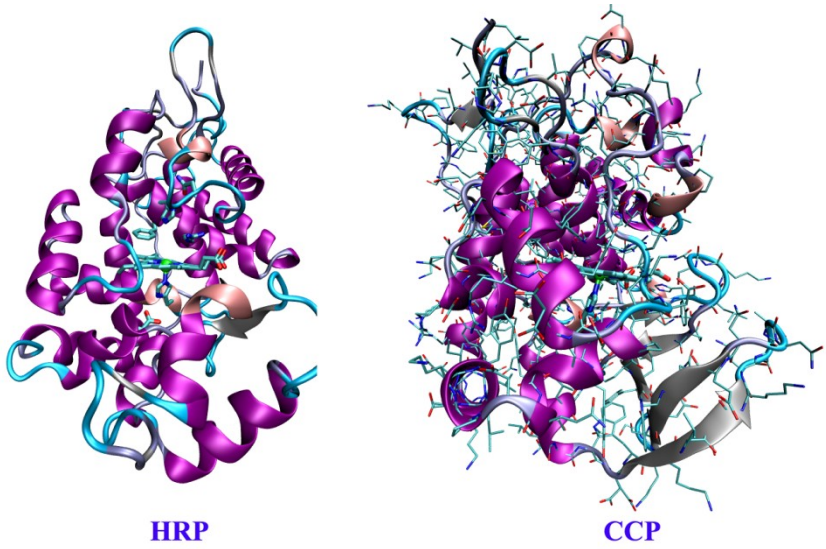


Fig. 8 crystal structure of HRP (pdb code 1ATJ)[38] and CCP (pdb code1ZBY).[62]

In the resting state, peroxidase contains an iron ion in the oxidation state +3, that is five coordinated to the four pyrrole nitrogens of the heme and to a nitrogen atom of an axial histidine, the proximal histidine (with the exception of chloroperoxidase which contains a thiolate ligand),[63] while the sixth coordination position is free. The catalytic process occurs through

several steps (**Fig. 9**). The first step is the heterolytic cleavage of the oxygen-oxygen bond of H_2O_2, followed by the release of a water molecule and concomitant two-electron oxidation of the protein, with the incorporation of one of the oxygen atom of H_2O_2 in an intermediate, compound I, that contains an oxyferryl (Fe(IV)=O) center and an organic cation radical which can be located either on the heme or on a protein residue, depending on the enzyme.

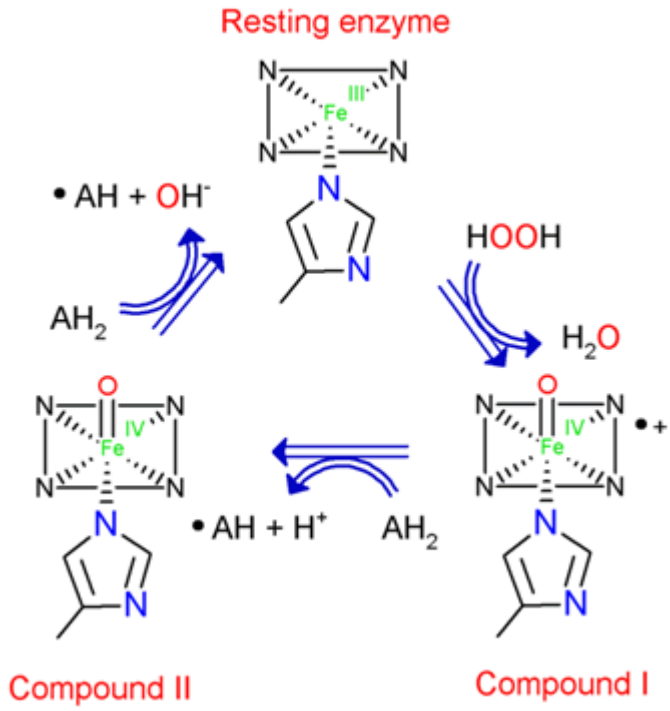


Fig. 9 Peroxidase catalytic cycle showing the enzyme intermediates, compunds I and II.

Compound I is able to oxidize a wide range of reducing substrates by a mechanism involving two single electron-transfers. First, it oxidizes one substrate molecule(s) to give a substrate radical and Compound II, in which the organic cation radical is discharged, while oxy-ferryl(FeIV=O) is still present (single electron transfer mechanism). Subsequently, compound II is

reduced by a second substrate molecule to the resting iron(III) state.[64,65] On the proximal site, the N_δ of the proximal histidine forms strong hydrogen bond with an aspartate side chain (Asp^{247} in Hrp, Asp^{235} in CCP) (**Fig. 10)** shows details of the proximal and distal site of HRP). This determine an increase in the basicity of the coordinating histidine, thus stabilizing the ferriheme form in the protein resting state (push effect). The binding of H_2O_2 occurs in the distal heme pocket, which is characterized by two highly conserved residues in all plant peroxidase: the distal histidine and the distal arginine, that are involved in peroxide activation and in the formation of compound I, being implied in the acid-base catalytic cleavage of the O-O bond (pull effect). The distal histidine is thought to promote the formation of Compound I, by acting as a general acid-base catalyst that deprotonates H_2O_2 facilitating the formation of a ferric peroxide (Fe-OOH) complex, and subsequently transfers the proton to the distal oxygen of the ferric peroxide complex, in the dioxygen bond cleavage step.

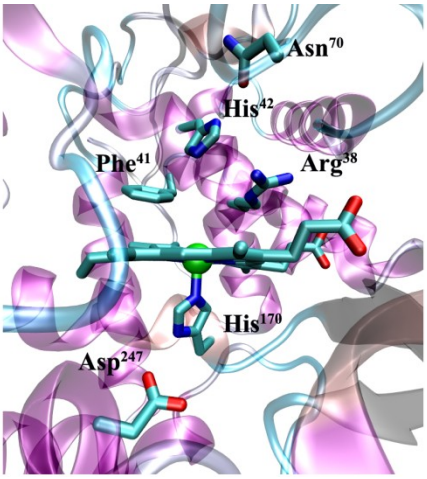


Fig. 10 Heme binding site of of

The conserved distal arginine (Arg^{38} in HRP, Arg^{48} in CCP) is involved in the charge stabilization, (mediated by its positively charged guanidinium group) of the developing OH- leaving group of hydrogen peroxide. Further, once the O–O bond is cleaved, the distal arginine contributes to stabilize the oxoferryl specie via hydrogen bond formation.

The important role of these highly conserved residues has been demonstated by site-directed mutagenesis, since their replacement causes a drastic reduction of the rate of compound I formation. [61,66-68]

1.6. Cytochromes

Cytochromes are cellular heme-containing protein whose principal function is electron transport. A cytochrome is defined as an heme-protein whose function involves transfer of reducing equivalents associated with a reversible change in oxidation state of the prosthetic group. Formally, this redox change involves a single-electron, reversible equilibrium between the Fe(II) and Fe(III) states of the central iron atom.

These molecules were initially described in 1885 by C.A. MacMunn as respiratory pigments isolated in tissues and organs of invertebrate and vertebrate animals, and subsequently identified as heme containing proteins in 1925 by David Keilin, that introduced the name "cytocrhome" and their current classification, based on the characteristic of their visible-light absorption spectra.[69,70] In fact, cytochromes have distinctive visible-light absorption spectra, and are classified depending on the position of their lowest energy absorption band in the reduced state, as cytochromes a (~605 nm), b (~565 nm), and c (550 nm). Within each class, early cytochromes are numbered consecutively, e.g. cyt c, cyt c1, and cyt c2, with more recent examples designated by their reduced state alpha-band maximum, e.g. cyt c559. An exception is cytochromes P450, which is named on the basis of the wavelength of the most intense absorption band (Soret band) of the carbon monoxide complex of the reduced form. Various amino acids are known to serve as axial ligands in cytochromes, also if bis-His and His-Met coordination (found as example in cyt b5 and cyt c554 (**Fig. 11**) and cyt c, respectively) are the most common coordination states. Other coordination states, such as bis-Met and His-Lys, have also been found. Interesting, ligation from the terminal α-amino group of a peptide bond has also been observed in cyt f.[55] Because they does not have a free coordination position

on the heme iron, bis-His and His-Met cytochromes are unable to activate exogenous molecules, such as oxygen and peroxide. Instead, they act as electron transfer component in many essential biological process, like respiration (cyt b, cytochrome c3),[71-73] photosynthesis (cyt b6),[74] desaturation of fatty acids (cyt b5),[75] nitrification pathway.[46] In these cytochromes, the iron ion is in the low-spin configuration in both the oxidation states.

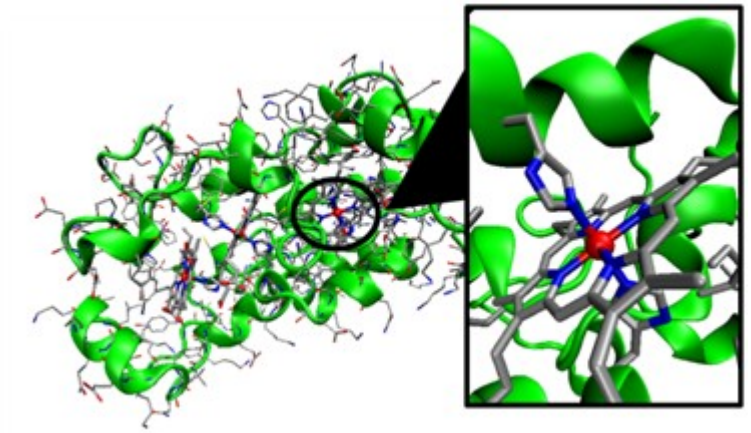


Fig. 11 Cytochrome C554.

Furthermore, the iron redox potential is finely modulated by the local protein environment surrounding the heme group. Overall, the redox potentials of all known cytochromes span a range of almost 800 mV, that corresponds, in thermodynamic terms, at about 17 kcal/mol in free energy.

Certainly, this variation is in part due to differences in the heme- iron first coordination shell, as well as to modification of the peripheral substituents on the porphyrin (in various heme forms, such as a, b, and c), that cause a shift of about 150 mV.[76] Clearly, there are additional factors that modulate the redox potential of the heme iron. In particular, the level of exposure to solvent, the heme binding mode, the charge distribution in the

heme environment, the orientation of the axial ligands. As an example, a large number of negatively charged residues around the heme favor a high oxidation state for the iron, while the progressive change from a negative charge distribution in cytochromes b to a positive one in cytochromes c, determines a gradual increase in the Fe(III)/Fe(II) redox potential. The efficient control of the reduction potential of the heme group is of crucial importance in electron transfer proteins, as in the biological electron transfer chains, the midpoint redox potential of each component dictates the direction of favorable transport between the individual constituents.

For this reasons, intense effort has been put into both experimental and theoretical analysis aimed at clarify the factors that determine the reduction potential of these metallo-proteins. However, due to the intricate set of factors that contributes to determine the redox potential, these studies often failed to identify a clear correlation between the reduction potential and any specified structural parameter (like heme solvent accessibility, number of protein-heme hydrogen bonds, nonbonded contacts).[43] Besides that, some correlations can be observed limiting the analysis to very closely related heme proteins, for example between the heme reduction potential and the solvent accessibility of the heme group.

1.7. Artificial metallo-proteins

Metallo-proteins are specially attractive targets for protein engineering, as protein design is an effective approach for mimicking metal coordination environments, and develop new systems that reproduce active site structures within a peptide scaffold.[77-81] These systems can be very useful in elucidating structural features that may remain hidden in site-directed mutagenesis studies, as the ability to reproduce desired active sites in a designed system is a true test of proper knowledge of protein folding and functions.[1] As design requires the incorporation of all the structural features needed to attain a function, the designed compounds should be of sufficient size and chemical diversity to allow the construction of functional sites. At the same time, they should be simple enough to avoid much of the ambiguity of interpretation associated with large proteins. Furthermore, reduced dimension respect to the native proteins, make these artificial metallo-proteins easier for practical applications, which spans from biocatalysis and biosensor technology, to degradation of pollutants or biomass, and drug or food processing.[82]

However, metallo-protein design has proven to be more challenging than the design of nonmetallo-proteins, since it requires a properly folded protein scaffold containing the correct number and type of ligands with the appropriate geometry to bind and activate the metal for chemical catalysis. On the other hand, most metal-binding sites are highly chromatic and display typical magnetic properties, allowing easier characterization by means of metal-based spectroscopic techniques rather than X-ray crystallography or nuclear magnetic resonance (NMR).[83,1] Since the first reports of artificial metallo-proteins appeared over 30 years ago, important progress has been made in reproducing many native metal sites, and several examples of

functional metallo-proteins equipped with useful electronic or optoelectronic properties, high enzymatic activities or new functions have been recently developed (**Fig. 12**). [77, 84,85]

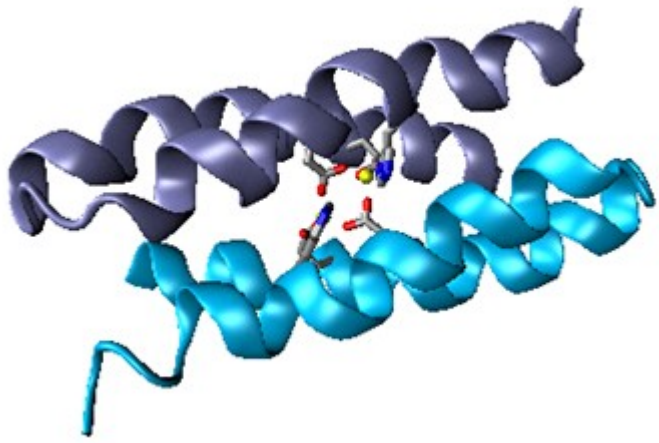


Fig 12. NMR structure of a de novo designed four-helical bundle di-iron protein with phenol oxidase activity.

Such compounds, equipped with high stability and great efficiency can be very useful as cheaper, more stable, and environmentally friendly novel chemical catalysts. Although some metal binding compounds have been designed through combinatorial and evolution methods (as example, there has been an attempt at creating new metallo-antibodies by affinity selection from a combinatorial library), most novel metal binding sites constructed in proteins have been designed using some form of rational, structure-based design.[86]

The rational design of metal binding protein can be divided into two stages: the design of the overall scaffold and the design of active sites.[87] However, the two design step are closely linked: the construction of metal binding sites requires the integration of scaffold and active site design, as amino acids should be placed in the appropriate geometry to form the metal ion coordination shell, while also maintaining steric compatibility with the

rest of the protein fold. Furthermore, in order to optimize metal-binding sites for function, active site charge neutrality, electrostatics and metal–ligand bond strain should also be considered.[88,89] As a result, the rational engineering of metallo-proteins is an iterative process of design, synthesis and rigorous characterization, until the desired properties are obtained. Because of the complexity of protein folding, where a large number of molecular interactions contributes to stabilization of the three dimensional structure and to the fine-tuning of the properties of metal cofactor, several redesign cycle are usually required, and new compounds are often obtained only through considerable trial and error. Despite this, three main structure-based design methods proved to be specially appealing for the engineering of novel metal binding sites into proteins:

- design or re-design of new metal-binding sites into native scaffolds;[90,91]
- minimalist approach, that focus on determination of the structural features necessary to confer the structure and function of the enzyme, and involves the construction of "minimalist" sequences that are simpler than their natural counterparts but, nevertheless, retain sufficient complexity for folding and function; [92]
- de novo design, that involves the construction of a protein, intended to fold into a precisely defined 3-dimensional structure, with a sequence that is not directly related to that of any natural protein.[93-99]

1.8. Artificial heme-proteins

The interest in engineering artificial heme-proteins as model complexes for natural proteins take advantage of great availability of structural data, as well as on potential application of heme-protein models. Also, the design of heme-protein synthetic analogues provides a mechanism by which to separate each factor that contribute to metal reactivity modulation in native heme-proteins, revealing key aspects on structure-function relationships. Indeed, great attention has been focused on the design of artificial heme binding proteins, and several artificial heme-proteins have been engineered, providing key insight on metal proprieties modulation by the protein matrix. The time-honored synthetic analogue approach to heme-proteins starts from the development of small metalloorganic complexes based on natural and synthetic metalloporphyrins with exogenous axial ligands, as catalysts for a wide range of industrial applications.[5] This approach focused on elucidating the effects of metal ion coordination on the spectroscopy, electrochemistry, and chemical reactivity of the heme group in the absence of the influence of the protein matrix. This traditional approach has further evolved in the development of more sophisticated molecules, that include small peptides as axial ligands, and are better models of native heme-proteins.[5]

As example, Benson et al. developed a class of artificial heme mimetics, named PSM, that are made up by two identical peptides covalently linked to iron(III) mesoporphyrin IX via amide formation between the heme propionyl groups and the amine of lysine side chains (**Fig. 13**). The peptide chains bear an histidine that acts as axial ligand to the heme iron, to form a bis-His complex.[100]

The peptide sequence was also designed such that coordination of the imidazole side chain of the histidine residue in each peptide to the heme metal would lead to helix induction. These models were extensively characterized, by means of CD, UV-vis and NMR. The spectroscopic characterization confirmed bis-His ligated state of the heme and significant helical content, that increased upon addition of alcohol co-solvents.[101,102]

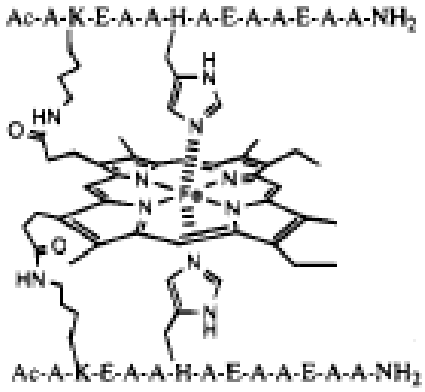


Fig.13 Reprinted with permission from Benson et al., J. Am. Chem. Soc., 1995. Copyright (1995) American Chemical Society.

More recently, the engineering of novel metallo-proteins is also progressing from purely structural to more functional designs, and new heme-proteins models with improved properties such as higher stability, greater efficiency, and specific functions are emerging.[103,104] As example, Koder et al. engineered an artificial oxygen transport protein, as a model of globin.[105] This model was completely de novo designed, beginning with a simple sequence composed by only three amino-acid (E,K,L), that assemble to form a four-helix-bundle motif. Then, two leucine residues at internal position were replaced with histidine, in order to anchor the heme cofactor through bis-histidines coordination. Afterwards, the sequence was adjusted to improve structural resolution and to facilitate O_2 binding. In particular,

water was excluded from the heme binding site by re-packing the protein interior and by changing the loop such that helical interface mobility was reduced, in order to avoid heme oxidation. The final model, shows O_2 affinities and exchange timescales that match natural globin with distal histidine. On the contrary, the binding of CO is weaker compared to that of natural globins, with the consequence that O_2 binds to this model tighter than CO (**Fig. 14**).

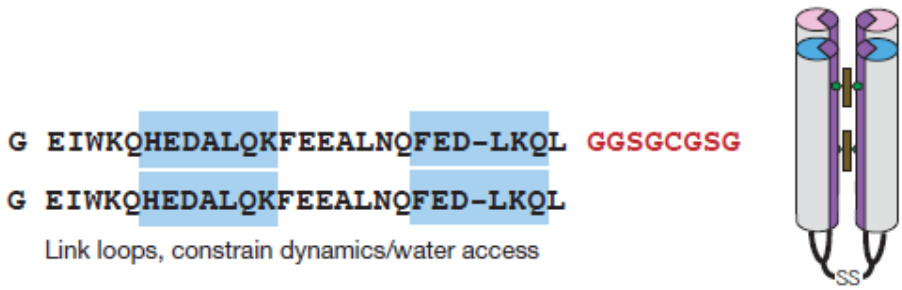


Fig. 14 Design of an artificial oxygen transport protein. Reprinted by permission from Macmillan Publishers Ltd, Nature, from Koder et al ref. 72., copyright 2009.

Many efforts have also been devoted to developing artificial heme compounds with peroxidase like activity, and some heme-protein models able of performing peroxidase activity with increase efficiency have appeared in literature. Microperoxidase (MPs) (**Fig. 15**), are a class of heme containing products derived from cyt c proteolysis, that retain the proximal histidine of the parent cytochrome as a ligand to the heme iron, and are characterized by mono-histidine-ligated state of the heme iron.[106-111] These molecules, have been very well characterized from different research groups, that obtain an impressive amount of information regarding their coordination properties, their secondary structure, and their aggregation process.

However, the structural characterization of MPs from a three dimensional point of view is made hard by their flexibility and oligomeric propensity. Also, the analysis is complicated by the solution speciation of the MPs in aqueous buffers. In fact, the spectroscopic titration as a function of pH indicated that there are several pKa observable in the MP8 and MP11 electronic spectrum, with more than one species existing in solution in the pH range 2-7.

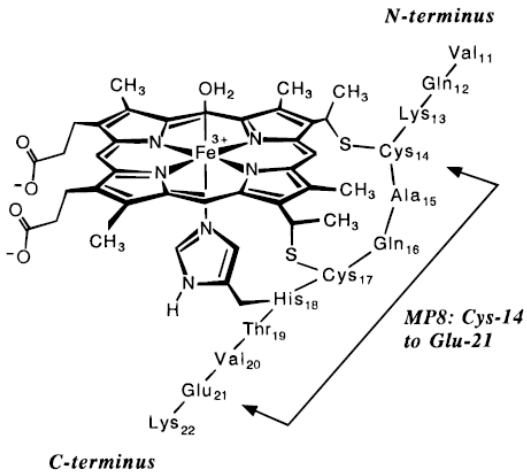


Fig. 15 Schematic representation of microperoxidase 8, depicted as six-coordinated aqua complex; the sequence of MP8. Reprinted with permission from Heme-Peptide Models for Hemoproteins. 1. Solution Chemistry of N-Acetylmicroperoxidase-8, Munro et al., Inorg. Chem, 35, 3752-3767. Copyright (1996) American Chemical Society.

An example of small artificial heme models equipped with peroxidase activity was provided in 1997 by Hecht et co-worker, that used the binary patterning approach in the design of semi-random libraries of peptides, and verified their ability to bound heme and to function as peroxidases. In the binary patterning approach, the patterning of polar and nonpolar amino acids is specified explicitly, but the exact identities of the side chains is varied extensively. Thus, a binary pattern designed to fold into a specific structure is compatible with many different amino acid sequences. 30 different binary

code sequences were synthesized, and 15 resulted to bind heme, with their absorption spectra and resonance Raman spectra resembling those of natural cytochromes.[112]

Although the binary code approach seemed well-suited for exploring the sequence requirements for designing particular structures, the combinatorial underpinnings of this strategy preclude the explicit design of catalytic sites. However, also if these sequences were not explicitly designed for function, several of these de novo heme proteins were able to catalyze the oxidation of TMB (2,2′,5,5′-tetramethyl-benzidine) and ABTS (2,2′-azino-di(3-ethyl-benzthiazoline-6-sulfonic acid) activity at levels substantially above background, with at least one protein that exhibits a rapid catalytic turnover. Subsequently, Hatch and co-worker constructed large libraries of synthetic genes encoding for four-helix bundle binary patterned proteins. Several of these proteins were found to bind cofactor and/or catalyze reactions. Remarkably, the NMR structure of two of these proteins was solved, confirming that these proteins form well-folded 4-helix bundles, with the hydrophobic residues hidden in the core and the hydrophilic residues exposed on the exterior.[113,114] Recently, these proteins have been used as starting point for mutagenesis studies, that were followed by color- based screening of peroxidase activity, showing in some case activity 3-fold higher than the parental sequence. As a consequence, these studies demonstrated that de novo proteins can be used as a novel feedstock for the evolution of enzyme activity.[115]

Also, one of the first attempts to rationalize the influence of the peptide sequence in the modulation of the heme properties is due to Casella and co-workers, that synthesized several synthetic heme porphyrin models made up by a ten residues peptide covalently bound to a porphyrin propionate

(through the amino terminal group), that were able to catalyze the oxidation of L- and D- tyrosines to o,o'-dityrosine in the presence of hydrogen peroxide. Also, the observed oxidation rate was different for the two enantiomers, thus highlighting a steroselectivity effect, which is due to the interaction with the chiral peptide chain. This effect was modulated by the composition of the peptide chain. In fact, reversed preference with respect to the oxidation of the two enantiomers was observed for the analyzed peptides. Starting from these conjugated, the authors added an L-histidine methyl ester residue to the other propionic group, in order to obtain a mono-his complex that reproduce the "push" effect of the proximal histidine, in which the distal plane of the porphyrin is covered by the deca peptide. They observed an improvement of the activity of the model systems, while different steroselectivity effects were still observed upon peptide composition.[116,109]

More recently, the artificial metallo enzyme group, where this project has been carried out, designed an artificial five-coordinated heme-complex with peroxidase activity, that combines the excellent structural properties of four-helix bundle protein scaffolds with the activity of natural peroxidases, that was named MP3 (MiniPeroxidase 3) (**Fig. 16**).[117]

MP3 is made up by two 41-residue helix-loop-helix (a2) motifs, assembled into an antiparallel four-helix bundle, with a covalently bound deuteroporphyrin housed in the bundle interior, to obtain an asymmetric helix– loop–helix/heme/helix–loop–helix sandwich arrangement, similar to the one found in bacterioferritin (an oligomeric protein containing both a binuclear iron center and a heme b). The heme environment design was based on the structural and functional properties of natural protein HRP. In particular, a histidine coordinating residue was introduced to reproduce the HRP proximal histidine, (which forms a hydrogen bond with an aspartic

side-chain, to improve the imidazolate character of the coordinating His, as in HRP proximal site), and an arginine residue was modeled in the distal site, mimicking Arg^{38} of HRP, at the right distance from the metal center to assist in hydrogen peroxide activation.

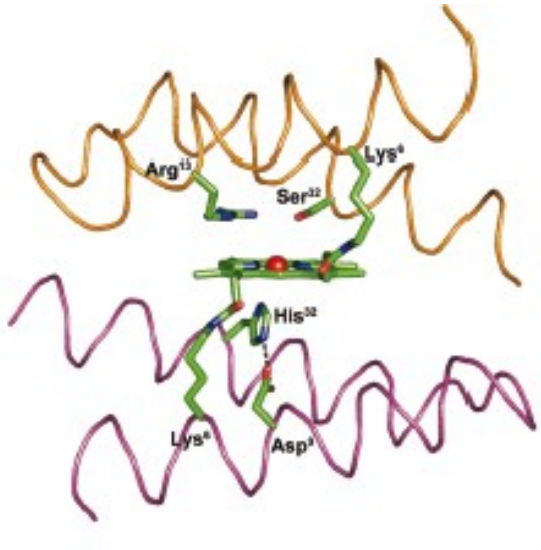


Fig. 16 Computer model of MP3. Reprinted from Chem. Eur. J., Vol. 18, pp. 15960 – 15971. Faiella et al., De Novo Design, Synthesis and Characterisation of MP3, A New Catalytic Four-Helix Bundle Hemeprotein, Copyright (2012), with permission from WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

MP3 was characterized by CD spectroscopy, that revealed high helix-forming propensity, while UV/Vis spectroscopy was used to analyze the coordination properties, confirming mono-His coordination above pH 4. In addition, kinetic experiment showed that Fe(III)-MP3 has turnover number (Kcat) only 8-fold lower than that of HRP, in the experimental conditions for maximal activity for each enzyme, while Fe(III)-MP3 Kcat is 10-fold higher than that of HRP at neutral pH.

1.9. Mimochromes

The *Artificial Metallo Enzymes Group* (AMEG) in Naples, where this work was carried out, has a long experience in mimicking metalloproteins in small peptide-based molecules for the selection of new advanced catalysts. A whole family of synthetic heme-protein models, named Mimochromes, was developed and fully characterized. These molecules are characterized by a sandwich like structure, in which the heme prosthetic group is embraced between two helical peptides, that are covalently linked to deuteroporphyrin IX (DPIX) through two amide bonds between the heme-propionic groups and the ε-amino groups of two lysine residues (**Fig. 17**).

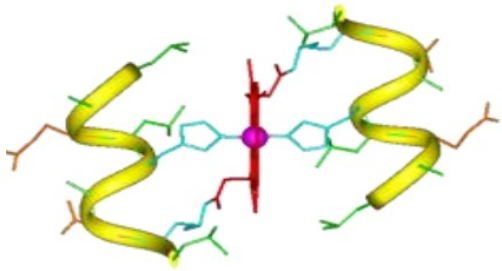


Fig. 17 Mimochrome I.

The prototype molecule, Mimochrome I, was designed through a miniaturization process, applied to the F helix of hemoglobin β-chain. This miniaturization process was based on detailed analysis of the target segment, in order to identify the constituents necessary for reconstruction of the target structure. This lead to identification of a nine residue segment (Leu^{88}- Leu^{96} in hemoglobin chain β) as the minimum requirement for complete coating of one face of the heme. A histidine residue (His^{92} in hemoglobin) that is located in the middle of this segment, binds the heme iron. Two leucine residues, at positions i-4 and i+4 relative to the histidine, hydrophobically

interact with the heme macrocycle and stabilize the sandwich structure. (Leu^{88} and Leu^{96} in hemoglobin) (**Fig. 18**).[118-120]

In order to stabilize the α helical conformation, some substitution were made to the target sequence:

- Ser ^{89} and Cys ^{93} have been replaced by two alanine residues, Ala^{2} and Ala^{6} (the CH_3 group of Ala experiences no loss in conformational entropy when a randomly coiled polypeptide folds to an α-helix)[121]
- Glu ^{90} and Asp^{94} have been replaced with glutamine and asparagine, (Gln^{3} and Asn^{7}, respectively) in order to remove charge that destabilize the desired conformation.

Hemoglobin Leu^{88}-Leu^{96} L^{88}-S^{89}-E^{90}-L^{91}-H^{92}-C^{93}-D^{94}-K^{95}-L^{96}

Mimochrome I Ac- L^{1} -A^{2} -Q^{3} -L^{4} -H^{5} -A^{6} -N^{7} -K^{8} -L^{9} - NH_2

Fig. 18 Comparison of Mimochrome I peptide sequence with the sequence of the F helix of hemoglobin β-chain target segment.

Finally, a lysine residue was introduced in position 8, in order to anchor the heme group. In fact, the peptide chain was covalently linked to the heme through amide bonds between the ε-amino groups of this Lys residue and a heme-propionic group. Moreover, two copies of the peptide were linked to each heme group, through covalent bond to the two porphyrin propionic groups, obtaining a pseudo-C2-symmetric dimer.

The spectroscopic and structural characterization of Mimochrome I indicated that it binds cobalt and iron in a low-spin bis-His-ligated state, as intended in the design. CD spectra in the far UV region confirmed the peptide chains to be predominantly in an α-helical conformation in both the apo and metalated species. Moreover, the insertion of metal ion into the porphyrin ring gave two diastereomeric forms, differing for the configuration

around the metal ion; that were separated as Co(III) complexes by RP-HPLC (**Fig. 19**).

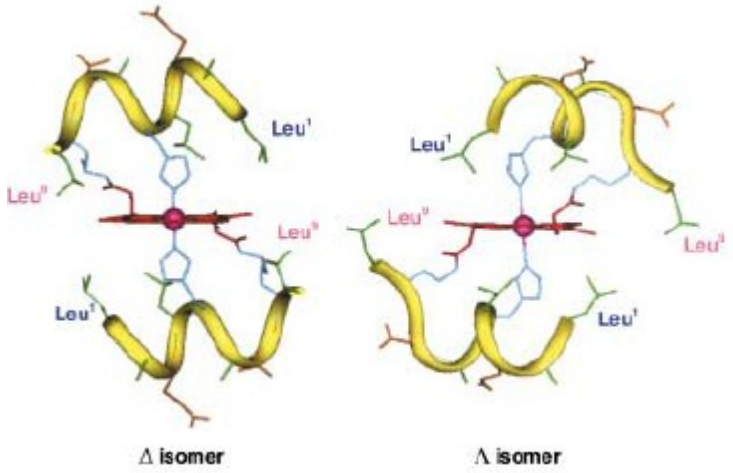


Fig. 19 average molecular structures of Co(III)-Mimochrome I Δ and Λ isomers, as obtained from NMR experimental data and RMD calculations.
Reprinted from Chem. Eur. J., Vol. 9, Lombardi et al., Design of a New Mimochrome with Unique Topology, pages 5643-5654., Copyright (2003), with permission from WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.

The information derived from Mimochrome I was used for improving the design, and develop new models with improved structural and functional properties. Two different strategy were followed:

- elongation of the peptide chains at the C-termini, with a four residue fragment, modeled in an extended conformation;[122]
- introduction of amino acidic residues able of forming specific intra-molecular, inter-chain interactions, that stabilize a single topology of the peptide chains around the porphyrin and favor helix formation.

In particular, the last strategy lead to the development of a re-designed molecule that adopts a unique topology: Mimochrome IV.[123,124] In Mimochrome IV, an intramolecular inter-chain Arg-Glu ion pair was

introduced, in order to favor the formation of only one diasteroisomer. In fact, Mimochrome IV differs from Mimochrome I for Glu^{1} and Arg^{9}, which replace Leu^{1} and Leu^{9}, respectively. Additionally, the solvent exposed Ala^{2} and Ala^{6} in Mimochrome I were replaced by Ser residues in Mimochrome IV, in order to further increase the water solubility of the new compound. The spectroscopic characterization of Mimochrome IV iron and cobalt complexes confirmed a bis-His axial coordination and a predominant helical conformation of the peptide chains. Moreover, it indicated the presence of the Λ isomer as unique specie. Subsequently, the Co(III) complex was structurally characterized both in solution and solid state, confirming the unique presence of the Λ topology (**Fig. 20**). Furthermore, the structural characterization evidenced that in the solid state crystal packing interactions strongly affects the molecular folding, with intra-chain Glu^{1}–Arg^{9} ion pairs that are preferred over the designed, and experimentally found in solution, inter-chain interactions.

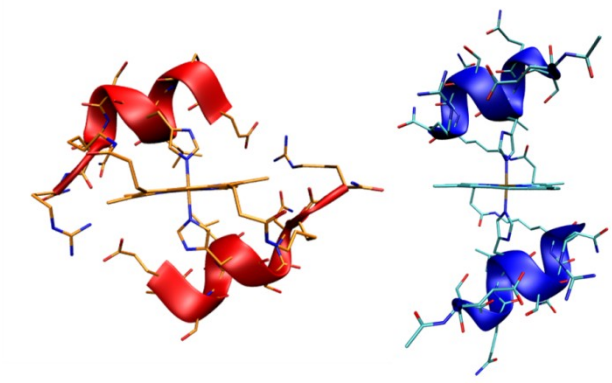


Fig. 20 NMR and X-ray structure of Mimochrome IV.

1.10. A heme–proteins model with peroxidase-like activity: Fe(III)-Mimochrome VI and its analogues

More recently, Mimochrome architecture was redesigned with the aim of introducing peroxidase like activity. In the new model, named Mimochrome VI (**Fig. 21**), the two chains have different length and composition: one is a 14-residue peptide chain, bearing a His residue for metal ion coordination, the other is a 10-residue peptide chain, that is devoid of a heme-coordinating residue, and therefore may create a cavity around the metal ion (distal cavity).[7] In this peptide, a Ser residue replace the coordinating histidine. Serine was chosen in order to facilitate the access of small ligands, due to its small side chain. Moreover, the serine side chain is also able to act as hydrogen bond donor, and may form hydrogen bond with exogenous molecules, stabilizing catalytic intermediates.

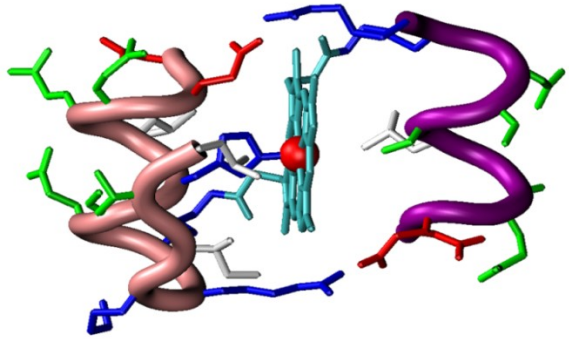


Fig. 21 Schematic representation of the model structure of Fe(III)–Mimochrome VI.

The two peptide chains embrace the metalloporphyrin, adopting a sandwich like structure, similar to that observed to other Mimochromes. In terms of secondary structure, the decapeptide was designed to adopt a helical conformation, whereas the tetradecapeptide was designed to fold in a short helical conformation (residues1–9), a loop (residues 10 and 11) and a short

β–strand (residues 12–14) that folds back (through the loop) to interact with the helical part. As in Mimochrome IV, inter-chain ion pairs between the carboxylate side chains of glutamate residues (Glu^2) on one helix and the guanidine groups of arginine (Arg^{10}) on the other helix were introduced in the model, in order to stabilize the tertiary structure. The positively charged Arg^{10} and the negatively charged Glu^2 at the C-terminal and N-terminal ends, respectively, with the opposite sign relative to the helix dipole, may also provide further stabilization of the secondary structure. Finally, several glutamines and a serine were introduced in the solvent exposed positions to promote water solubility.

The spectroscopic and functional characterization indicated that Fe(III)-Mimochrome VI is an efficient heme protein model, which displays a peroxidase-like catalytic activity. By the use of hydrogen peroxide, it efficiently catalyzes the oxidation of several substrates, with a typical Michaelis-Menten mechanism and with several multiple turnovers. Fe(III)-Mimochrome VI peptide framework, despite its small structure (a total of 24 amino acid residues), confers higher efficiency to the porphyrin cofactor. Three important outcomes deserve highlighting: *i)* Fe(III)-Mimochrome VI efficiently catalyzes the oxidation of different substrates, such as ABTS and guaiacol, by activating H_2O_2, and efficiently catalyzes the nitration of phenols; *ii)* Fe(III)-Mimochrome VI displays a very high specific activity (104 mol mg^{-1} s^{-1} for ABTS oxidation), with respect to highly purified HRPs (91 mol mg^{-1} s^{-1} for ABTS oxidation at pH 4.6); these value highlights both its high catalytic efficiency and small molecular mass, compared to natural peroxidases that contain more than 300 amino acid residues; *iii*) Fe(III)-Mimochrome VI exhibits multiple turnover kinetics: more than 4000 turnovers within 10 min were observed in the ABTS oxidation. All these

features indicate that Fe(III)-Mimochrome VI is an attractive, low-molecular weight heme enzyme, which will serve as an excellent scaffold to further develop designed heme-based biocatalysts.

In order to achieve deep insight on the role of the peptide chains in modulating the structural and functional properties of Mimochrome VI, some point amino acid substitutions in its peptide chains were introduced. In particular, amino acid replacement were made into strategic position, with the aim of clarify the role of selected residues on the catalytic properties. In details, the following substitutions were introduced into Mimochrome VI:

- ✓ $Glu^{2}Leu$ in the tetradecapeptide and decapeptide chain;
- ✓ $Arg^{10}Leu$ in the tetradecapeptide and decapeptide chain;
- ✓ $Ser^{6}Gly$ in the decapeptide chain.
- ✓

These point substitutions were aimed to:

- ✓ break the inter-chain ion pair between the charged residues, thus allowing to understand the role of these interactions and, as a consequence, of the involved amino acids in the properties of the molecule;
- ✓ verify if the serine residue at the distal site of the heme is involved in the binding of substrates in the heme pocket.

Five different analogs of Mimochrome VI were synthesized and characterized (Rosa Vitale PhD. Thesis, 2013 and Vitale *et al*., manuscript in preparation).

Their ability to oxidize organic substrates (ABTS) was explored, and compared with those of Mimochrome VI and HRP, (table 2).[7,125] One of this

analogues, characterized by the replacement of Glu^2 on the tetradecapeptide with leucine, showed catalytic performance approaching that of HRP. In fact, the k_{cat} value observed for this minienzyme is only 7.9 -fold lower than HRP, in the experimental conditions for maximal activity for each enzyme, whereas the k_{cat} value of Fe(III)-Mimochrome VI Glu^2Leu TD is about 15 -fold higher than HRP at neutral pH. Moreover, Fe(III)-Mimochrome VI Glu^2Leu TD shows a significant improvement in the activation of H_2O_2, respect to Fe(III)-Mimochrome VI. In fact, the catalytic efficiency of the new developed heme-enzyme for the reducing substrate is about 3- and 4-fold higher than HRP and Fe(III)-Mimochrome VI, respectively. The catalytic activity of Fe(III)-Mimochrome VI E^2L TD, which is comparable to those of complex proteins, demonstrated that its miniaturized structure holds essential elements for tuning the iron-porphyrin peroxidase activity.

Table 2 Steady state kinetic parameters for H_2O_2 oxidation of ABTS

Enzyme	pH	Km ^{H_2O_2} (mM)	Km ^{AH_2} (mM)	Kcat (s^{-1})	Kcat/Km ^{H_2O_2} ($mM^{-1}s^{-1}$)	Kcat/Km ^{AH_2} ($mM^{-1}s^{-1}$)	Mw(E) (Da)	Specific Activity (mmlo $g^{-1}s^{-1}$)	TON *1e^{-3} [ABTS]/[E_0]
mimochrome VI[a]	6.5	44.0	0.0840	371	8.43	4417	3.551	104	4.00
E^2L TD[b]	6.5	31.3	0.0500	785	25.1	15704	3538	222	5.88
R^{10}L TD[b]	6.5	54.0	0.0380	685	12.7	18026	3508	195	5.56
S^6G D[b]	6.5	45.6	0.125	468	10.3	3743	3520	133	3.33
R^{10}L D[b]	6.5	19.7	0.0290	169	8.58	5828	3508	48.2	3.33
E^2L D[b]	6.5	95.6	0.115	381	3.99	3316	3538	108	3.57
HRP[b]	4.6	0.854	1.07	6209	7270	5803	44174	141	50.0
HRP[d]	7.0	0.0115	5.10	52.5	4565	10.3	44174	1.19	0.333

Adapted from Rosa Vitale PhD thesis.

[a] Data from reference 93; [b] data from Vitale PhD thesis, [c] data from reference 125.

1.11. Aim of the thesis.

The aim of this PhD project was to engineer new heme-protein mimetics that couple high activity with a well-defined three-dimensional structure, thus allowing structural characterization. As outlined in the previous paragraph, interesting results have been obtained on the heme-protein models belonging to the Mimochrome family. By using a re-design procedure and computational techniques, the catalytic performance of Fe(III)-Mimochrome VI were improved. The introduction of point amino acid substitutions revealed a key role of some residues in the modulation of the heme-model properties. One main drawback of Mimochrome VI and its analogues relies in the lack of detailed information of their three-dimensional structures.

Detailed structural characterization of Mimochrome molecules was straightforward for the *bis*-His–ligated complexes. In fact, as described in the previous paragraphs, solution structures by NMR were obtained for the fully diamagnetic compound Co(III)-Mimochrome I Δ and Λ isomers. Further, for Co(III)-Mimochrome IV, the structures in solution and in the solid state were solved.

In the cases of the *mono*-His ligated complexes such as Mimochrome VI attempts to solve the structure on the fully diamagnetic Co(III) complex were unsuccessful. In fact, analysis of the Nuclear Magnetic Resonance (NMR) spectrum of Co(III)-Mimochrome VI revealed the presence of different species in solution (**Fig. 22**). The complexity of the ^{1}H NMR spectrum, characterized by the presence of several sets of signals for both the porphyrin *meso* and pyrrole protons (**Fig. 22 a**) as well as for the coordinating His (**Fig. 22 b**) suggests the presence of different porphyrin

systems in solution, which precluded the structure determination. Moreover, effort to crystallize this molecule failed.

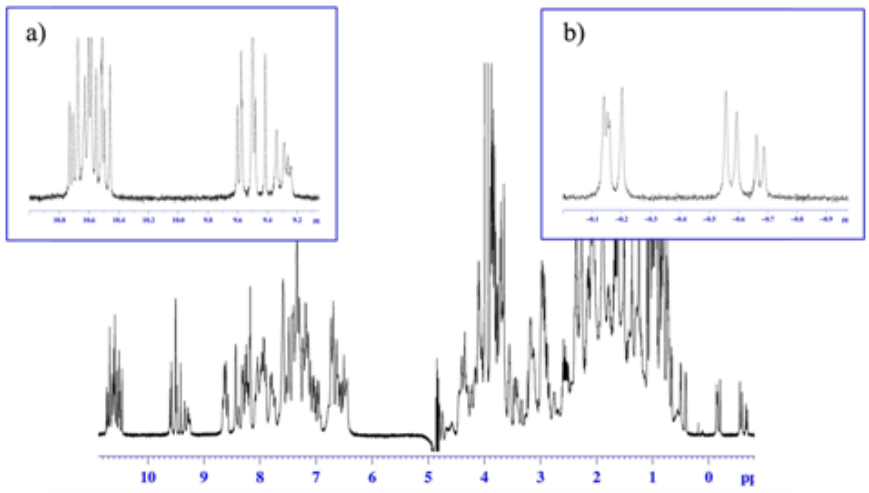


Fig. 22 1H NMR spectrum of Co(III)– Mimochrome VI.

Therefore, even though the catalytic results provide direct evidences that the miniaturized structure of Mimochrome VI and its analogues hold essential elements for a fine tuning of the iron-porphyrin peroxidase activity, a detailed structure-activity relationship analysis cannot be drawn. Thus, a very important goal in the development of bio-mimetic heme-containing catalysts is to obtain functionally active molecules with a well defined structure. This represents an intricate problem since the strategies used for the molecular design should take into account all the variables needed to specify both the structure and the function.

In this PhD project, two different strategies have been followed in order to overcome the structural flexibility of the parent Mimochrome VI molecule and design new mimetics:

- incorporation of non-coded amino acid with α-helix stabilizing effect;

- elongation of the peptide chains, that allows the introduction of a large number of intra-chain and inter-chain interactions.

The first strategy was based on the inclusion into Mimochrome VI of non coded amino acids, with a well known helix stabilizing effect. In this strategy, the new molecule is intended to share with the parent Mimochrome VI the same global framework, being characterized by the same secondary structural motifs. For this reason, the strategy dose not involved an ex-novo design of the protein scaffold, and does not required an extensive re-design process. The designed Mimochrome VI –2U1L was synthesized and the analysis of its Co(III)-complex structural properties carried out by ^{1}H-NMR spectroscopy.

The second strategy, involved the design of a new protein framework, characterized by a different organization of the secondary structure elements. In particular, the peptide chains were elongated in order to include a greater number of intra-chain and inter-chain interactions, aimed to stabilize the global fold. This strategy required an extensive re-design process, that involved several cycle of re-design and energy minimization. To this aim, the X-ray structure of a Co(III)-Mimochrome IV was used as template. In order to simplify the design and characterization, in this thesis a symmetric model, named Mimochrome VII, characterized by a bis-histidine metal binding site was obtained. The NMR structure of this symmetrical model has been used as a framework to design an asymmetrical penta-coordinated model. In particular, in Mimochrome VII was created a distal cavity that contains an arginine residue, mimicking Arg^{38} of Horseradish Peroxidase (HRP), and a non-

coordinating His residue to mimic His[42] of HRP. The structural characterization of this new model is under course.

Results and discussion

2. Development of a heme-protein model that incorporate α-aminoisobutyric acid (Aib), a non-coded amino acid.

2.1. Mimochrome VI-2U1L

The first strategy toward the engineering of new heme-protein mimetics was devoted to introduce into Mimochrome VI some structural constraints aimed at stabilizing the structure while preserving functionality.

As previously described, Mimochromes are composed of two peptide chains that surround the heme group and have the following two main features: i) the peptide chains are in α-helical conformation; and ii) the helix -heme-helix sandwich has a covalent structure. The use of α-helical peptides able to cage the heme group, has been and still remains the focus of the majority of the works in the field of heme-protein mimetics, because the α-helix motif is a recurring structural motif that surround the heme cofactor in numerous natural hemeproteins. Herein, α-aminoisobutyric acid, Aib (U), a well-established helix stabilizing non-coded amino acid that occurs extensively in peptides produced by microbial sources (for example in the antibiotic peptides chlamydocin, alamethicin, antiamoebin, emerimicin, and zervamicin), has been introduced in Fe(III)-Mimochrome VI, in order to favour the formation of the helical structure.

In fact, the replacement of the proton on the C^{α} atom in an alanine residue with another methyl group severely restricts the possible rotations about the N-Cα and C_{α}-C' bond, forcing the Phi and Psi values to fall in two very restricted regions near -57°, -47 ° and +57°; +47°, corresponding to α-helix or 3_{10}-helix structures (**Fig. 23**). Aib oligomers form exclusively 3_{10}-helices, and the strong propensity to form 3_{10}-helices is maintained when the Aib composition exceeds 50%. On the contrary, when the Aib content is less than 50%, mixed α/3_{10}-helical structures can occur, and several factors such

as length, composition and exact sequence, favor one or the other helix type, with peptide longer more than 6 residues (between 6 and 20 residues) that generally fold into predominantly α-helices.[126,127]

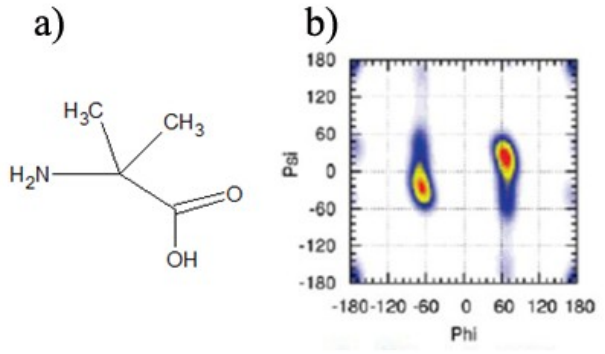


Fig. 23 a) α-aminoisobutyric acid, Aib (U); b) Ramachandran chart of Aib residue.

In this work, two Aib residue were introduced in Mimochrome basic structure, in order to favor helix formation and overcome the structural mobility that precluded the structural characterization of the original model. In particular, these substitutions were done on the most active analogues of Mimochrome VI, Fe(III)-Mimochrome VI E^2L TD.

This new mimetic, as the parent molecule, exhibits a mono-histidine five coordination to the heme, and it is made up of two peptide chains of different length and composition covalently linked to deuteroporphyrin IX via amide bond between the ε-amino groups of a lysine residue and the heme-propionic groups

The new model, Mimochrome VI-2U1L (**Fig. 24**), shares with the original template, Mimochrome VI, the following features:

- one chain is a decapeptide intended to fold in α helical conformation;
- one chain is a tetradecapeptide intended to fold as follows: α-helix (1-9), loop (10-11), β-strand (12-14). This peptide chain bears an histidine that acts as an axial ligand to the heme central ion.
- Glu^2 on tetradecapeptide was replaced with Leu. This substitution removes the Arg^{10}-Glu^2 ion pair interaction, thus enabling Arg^{10} to participate in the catalytic cycle, mimicking the Arg^{38} of the HRP cycle

Respect to the template Mimochrome VI E^2L TD, Gln^3 and Ser^7 on the decapeptide were replaced with two Aib residues, with the aim of stabilizing the helix, and overcome conformational mobility.

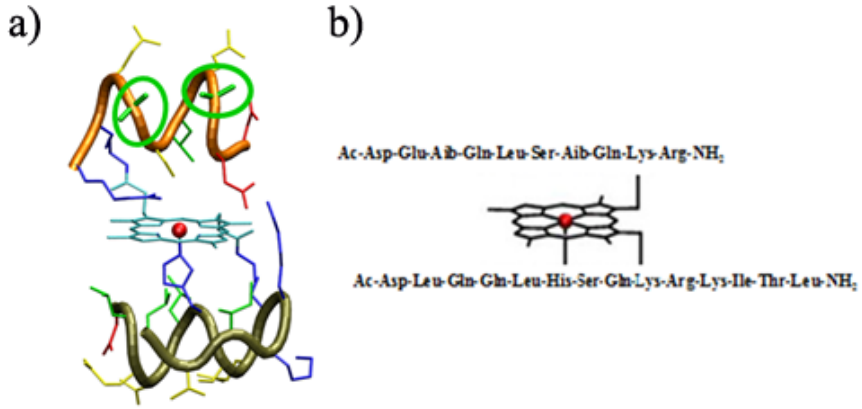


Fig. 24 Mimochrome VI-2U1L a) three-dimensional model; b) primary sequence.

2.2. Synthetic strategy: Mimochrome VI-2U1L apo form

Mimochrome VI-2U1L was synthesized as previously described for Mimochrome VI.[7] Peptide chains were synthesized in solid phase, by using 9-fluorenylmethoxycarbonyl (Fmoc) protection strategy. The selective deprotection from (4-methoxyphenyl)diphenylmethyl group (Mmt) of Lys^9 residues on the two chains is required before their cleavage from the resin. In this way, it is possible to obtain the peptides with the free Lys^9 and all other residue side-chains still protected. The peptides were then coupled to deuteroporphyrin IX in solution to afford, after side chain deprotection, Mimochrome VI-2U1L.

The synthetic protocol is made up of three main steps (**Fig. 25**):

- synthesis of the monosubstituted intermediate (decapeptide deuteroporphyrin-IX). Care was taken to minimize the formation of bis-substituted product by adding dropwise the solution containing the peptide to a large excess of DPIX.
- synthesis of the disubstituted intermediate (decapeptide-DEUIX-tetradecapeptide);
- deprotection of the disubstituted intermediate to obtain the *apo* form, Mimochrome VI-2U1L.

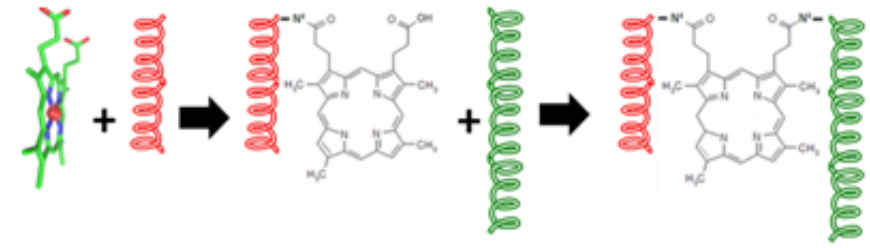


Fig. 25 Schematic representation of the synthetic strategy.

The monosubstituted intermediate was analyzed by RP-HPLC, on a C-8 column, using a linear gradient elution of water/acetonitrile 0.1% TFA.

The chromatographic profile, shown in **Fig. 26**, evidences the presence of two peaks of similar intensity at R_t of 26.97 and 27.23 minutes, respectively.

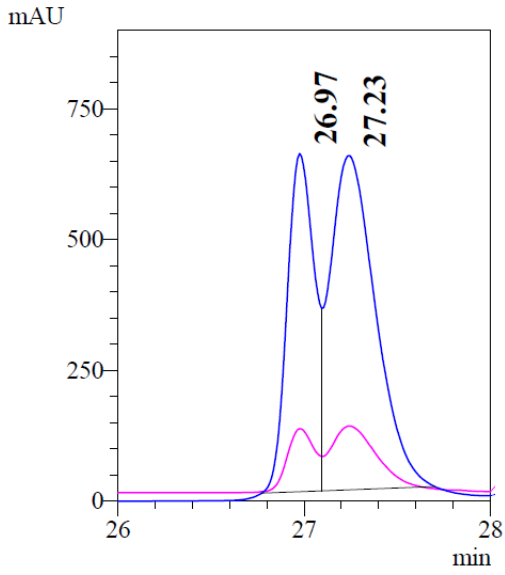


Fig. 26. Mimochrome VI-2U1L apo form chromatogram.

Further, the mass spectra of the two peaks showed similar behaviors: an intense peak at m/z 1147.0 ($[M-3H]^{3+}/3$), a second peak at m/z 860.5 ($[M-4H]^{4+}/4$) and a third peak at 866.6 ($[M-5H]^{5+}/5$). These values are in agreement for both peaks with the mass expected for Mimochrome VI-2U1L (3437.8 Da).

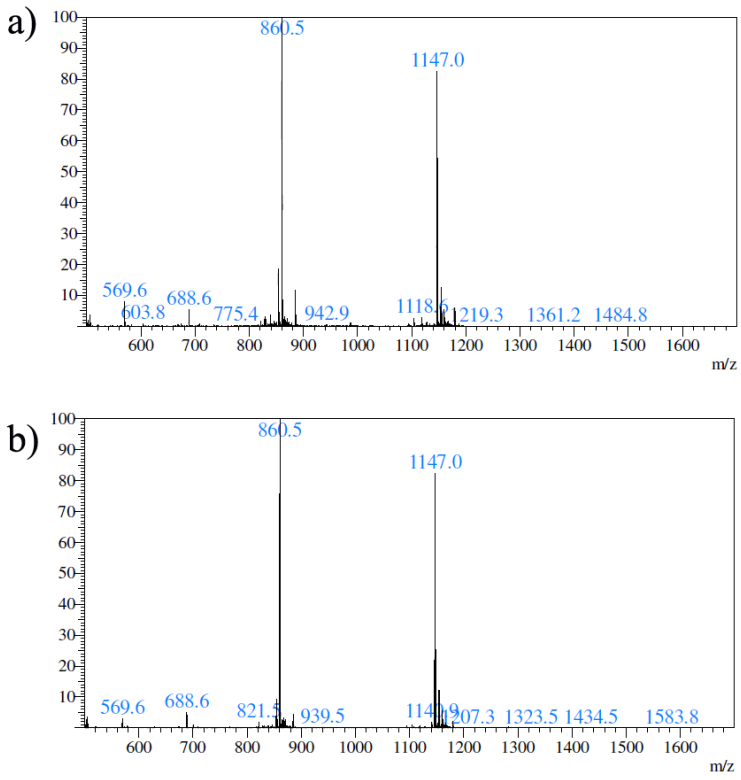


Fig. 27. Mimochrome VI-2U1L apo form ESI-MS spectra a) peak at R_t of 26.97 min; b) peak at 27.23 min

Therefore, we hypothesized that the apo form is actually constituted by an equimolar mixture of two constitutional isomers. In fact, in the first step of the synthetic protocol, coupling of the decapeptide to the deuteroporphyrin IX can occur at either the 2 or the 18 propionic group (see **Fig. 28**). As a consequence, the monosubstituted intermediate decapeptide deuteroporphyrin-IX is an equimolar mixture of two constitutional isomers mono-substituted at position 2 or 18 of the porphyrin ring. The subsequent coupling of the tetradeca-peptide to the monosubstituted intermediate decapeptide deuteroporphyrin-IX give an equimolar mixture of two Mimochrome VI-2U1L constitutional isomer (**see Fig. 28**). The presence of

the two Aib residues in this molecule, respect to the parent Mimochrome VI E^2L TD, may induce slightly different retention times for the two specie, thus enabling their separation. In fact, even though the synthetic protocol used for Mimochrome VI and its analogues can also afford to an equimolar mixtures of the 2(18)- decapeptide, 18(2)-tetradecapeptide constitutional isomers, their separation was not detectable by RP-HPLC.

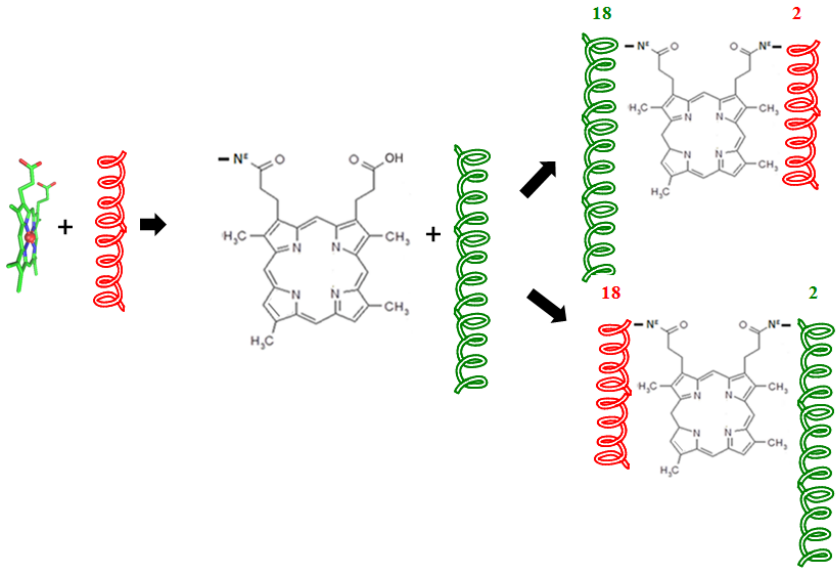


Fig. 28 Constitutional isomers of the Mimochrome VI-2U1L apo form.

2.3. Co(III)-Mimochrome VI-2U1L

With the aim of structurally characterize the new compound, the diamagnetic Co(III) ion was inserted into Mimochrome VI-U1L apo form mixture. Cobalt (III) was inserted into the molecule using Co(II) acetate according to the literature method.[45] **Fig. 29** shows the HPLC chromatogram of the reaction mixture. Four peaks, named peak a, peak b, peak c and peak d, were detected. No substantial changes in the ratio between the peak areas of these species occurred either by extending the reaction time or by favoring Co(II)→Co(III) air oxidation.

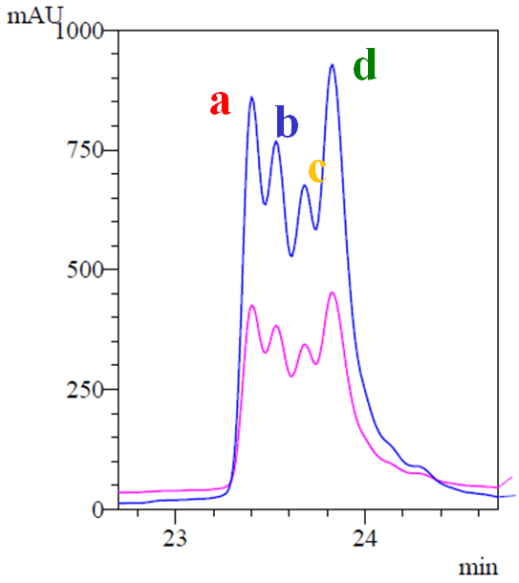


Fig. 29 RP-HPLC chromatogram of Co(III)- Mimochrome VI-2U1L.

The four species were then separated by preparative HPLC and separately analyzed (**Fig. 30**). The ESI-Mass spectra for all the peaks revealed a molecular weight of 3494.9 Da in agreement with the mass expected for Co(III)-mimochrome VI-2U1L. The UV-visible spectra of the four peaks showed similar behaviors, both for the intensities and positions of the Soret,

β and α bands. The UV-visible spectral data are in agreement with the presence of a Co(III) porphyrin complex characterized by a mono-histidine axial coordination, for all the four peaks.

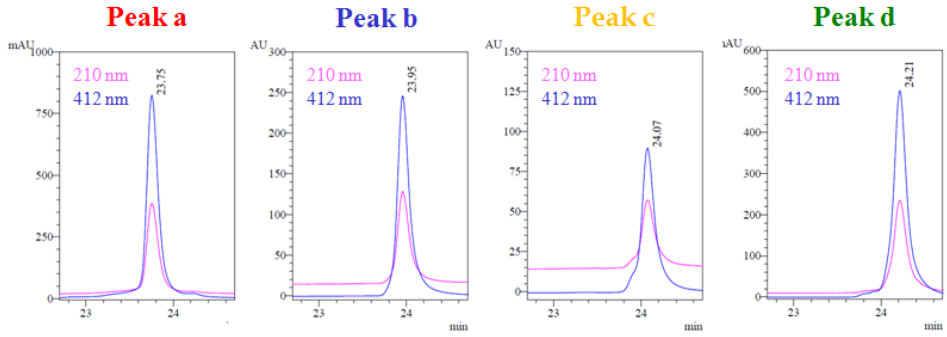


Fig. 30. RP-HPLC chromatogram of the four Co(III)-Mimochrome VI-2U1L species after purification.

2.4. Co(III) -Mimochrome VI-2U1L: NMR spectroscopy

The four Co(III)-Mimochrome VI-2U1L species, purified by preparative RP-HPLC, were separately analyzed by Nuclear Magnetic Resonance.

The analysis was carried out in H_2O/TFE 60/40 (v/v). The ^1H spectra are shown showe a single set of resonances for the deuteroporphyrin protons, and two sets of resonances for the peptidic part. The intensities of the deuteroporphyrin proton resonances as compared with those of the peptide chains are in the expected 1:2 ratio.

A careful examination of the NMR spectra revealed that a complete separation of the four species was not obtained. **Fig. 31** shows the high field region of ^1H NMR spectra, where resonate the His imidazole protons.

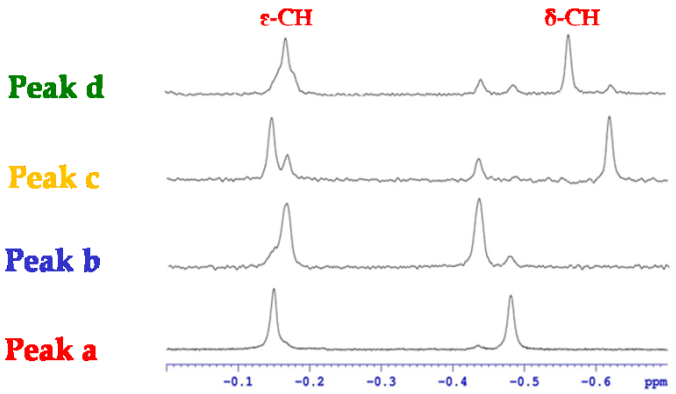


Fig. 31 High field region of Co(III)-Mimochrome VI-2U1L species 1D ^{1}H NMR spectra acquire in water/TFE 60/40 (v/v). The signals showed are attributed to the His ε-CH and δ-CH imidazole protons, that experience an extremely large ring current effect and appear in the high-field region of the spectrum.

In particular, while **peak a** seems to contain only one species, the NMR spectrum of **peak b** displays a predominant species with a small contamination (about 5%) of the species deriving from **peak a**. On the contrary, **peak c** and **peak d** show more complex spectra, due to the higher degree of contamination and to the presence of more than two species in solution. For this reason, a complete resonance assignment and structural characterization was achieved only for the first two peaks.

2.5. Co(III)-Mimochrome VI-2U1L resonance assignment

Peak a

Resonances assignment was accomplished by 2D experiments (NOESY,[128] TOCSY,[129] DQF-COSY[130]) by means of the sequential technique for both deuteroporphyrin and peptide protons.

Histidine sidechain protons were unequivocally assigned. The δ-CH and ε-CH imidazole protons experience an extremely large ring current effect and appear in the high-field region of the spectrum. The δ-CH resonances were easily identified on the basis of medium range NOE with the His^6 ββ'-CH_2 protons.

In the TOCSY experiment the δ-CH signal at δ = -0.49 is correlated to ε-CH imidazole proton at δ = -0.16 ppm. This last resonance is in strong dipolar contact with a resonance at δ = 9.23 ppm, which was assigned to the imidazole δ-NH of the histidine at position 6 of the tetradecapeptide chain.

All deuteroporphyrin proton resonances were unambiguously assigned, including those of the propionic groups, that exhibited clearly distinct resonances. In particular, the deuteroporphyrin assignment was easily performed following the NOE connectivities around the ring and starting from the 5H proton (at δ =10.50); it is the only proton in dipolar contact with the two methyl groups (3-CH_3 and 7-CH_3) (see **Fig. 32)** for deuteroporphyrin numbering).

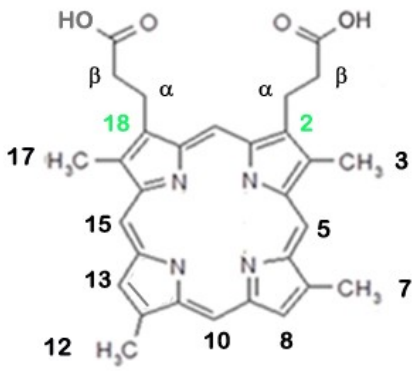


Fig. 32 Deuteroporphyrin IX.

Chemical shifts of peptide chains show αCH resonances significantly up-field shifted, relative to their random coil values (see **Fig. 33**),[131,132] and it is also shown by amidic NH, as well as by sidechain protons (data not shown). It is reasonable to assume that these deviations are due to the peptide secondary structures.[132-134] Furthermore, this effect is most pronounced for the tetradecapeptide chain (bearing the coordinating histidine), thus we could suppose that this peptide chain experiences deuteroporphyrin ring current too.

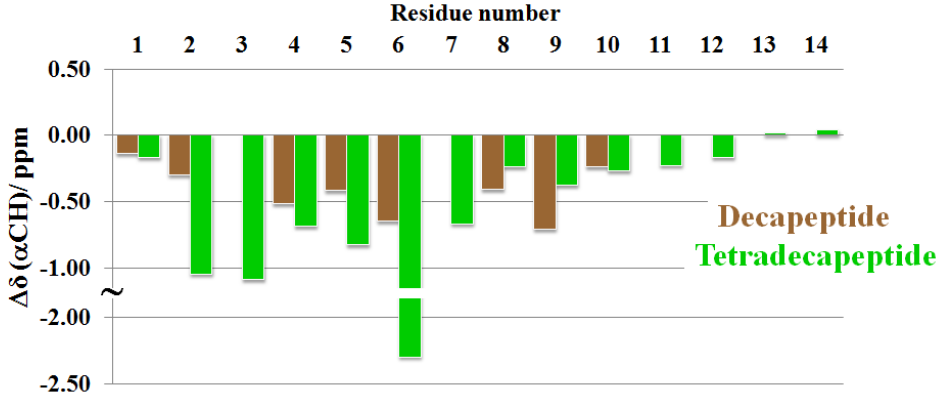


Fig 33 Chemical shift index $\Delta\delta_{\alpha CH} = (\delta_{osb} - \delta_{randomcoil})$ of **peak a**.

Indeed, the pattern of the NOE connectivities strongly supports the presence of a quite regular structure for the tetradecapeptide chain.

A total of 395 NOE connectivities were measured. The relative intensities of the structurally considerable NOE crosspeaks are shown in **Fig. 34**.

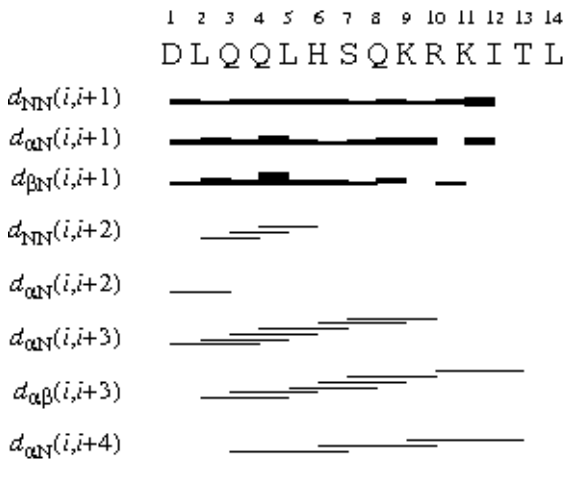


Fig. 34 Co(III)-Mimochrome VI-2U1L **peak a**:summary of NOE contacts. Bar thickness is proportionally related to the NOE intensity.

A continuous stretch of $d_{Ni\text{-}Ni+1}$ were monitored. Furthermore, several medium- and long-range contacts were unambiguously observed throughout the peptide sequence such as relatively strong $d_{\alpha i\text{-}\beta i+3}$, $d_{\alpha i\text{-}Ni+3}$, and weaker $d_{\alpha i\text{-}Ni+4}$ NOEs.

These data are consistent with an α-helix comprising residues 1-10, while it is possible to hypothesize a less regular structure n the C-terminal region (Lys^{11}-Leu^{14}).

From a structural point of view, the determination of the His sidechain orientation with respect to the peptide backbone is of interest. Several NOE contacts of the histidine protons with Leu^{2} and Leu^{5} and Lys^{9} protons were observed. Some of these contacts are unambiguous since they are caused by

the ε-CH and δ-CH protons. Unexchangeable δ-NH is also clearly visible, indicating that this proton is buried in the interior of the molecule and/or hydrogen-bonded. This further confirms that the cobalt(III) axial ligation occurs through the unprotonated imidazole Nε atom of the histidine, as already reported for other Mimochrome Co(III) complexes.[119,124]

Furthermore, the δ-CH and ε-CH imidazole protons have different connectivities with the deuteroporphyrin protons. These connectivities were crucial to define the orientation of the His^6 imidazole ring with respect to the deuteroporphyrin ring. In particular the following dipolar contacts are observed: εCH ↔ 8H, εCH ↔ 10H, εCH ↔ 12 CH_3, εCH ↔ 13H, εCH ↔ 15H, δCH ↔ 20H, δCH ↔ 3CH_3, δCH ↔ 5H. These connectivities indicate that the imidazole plane intersects the deuteroporphyrin ring through the 4 and 14 positions of the deuteroporphyrin ring (figure nnn).

Finally, it was easy to identify the tetradecapeptide chain as the one bound to propionyl group in position 2 on the basis of several NOE contacts between Lys^9 sidechain and propionic α,α'- and β,β'-CH_2 protons (**Fig. 35**)

Regarding the decapeptide chain, as already reported, even if the αCH chemical shifts are up-field shifted respect their typical random coil value, and indicative of an helical conformation, the NOEs data are not completely unambiguous due to severe overlaps. In particular, a continuous stretch of $d_{Ni\text{-}Ni+1}$ NOEs along the entire peptide sequence, and some medium-range contacts such as $d_{\alpha i\text{-}\beta i+3}$, $d_{Ni\text{-}Ni+2}$ were monitored and corroborated the existence of an organized structure of the peptide chain. Thus, at this stage, it was not possible to perform a complete structural analysis on the decapeptide chain. Different experimental conditions, such as temperature, pH, solvent, could help to overcome overlapping phenomena.

As for the tetradecapeptide chain, the NOE contacts between Lys^9 sidechain and propionic α,α'- and β,β'-CH_2 protons, allowed us to establish that this chain was linked to the propyonil group in position 18 of the deuteroporphyrin ring.

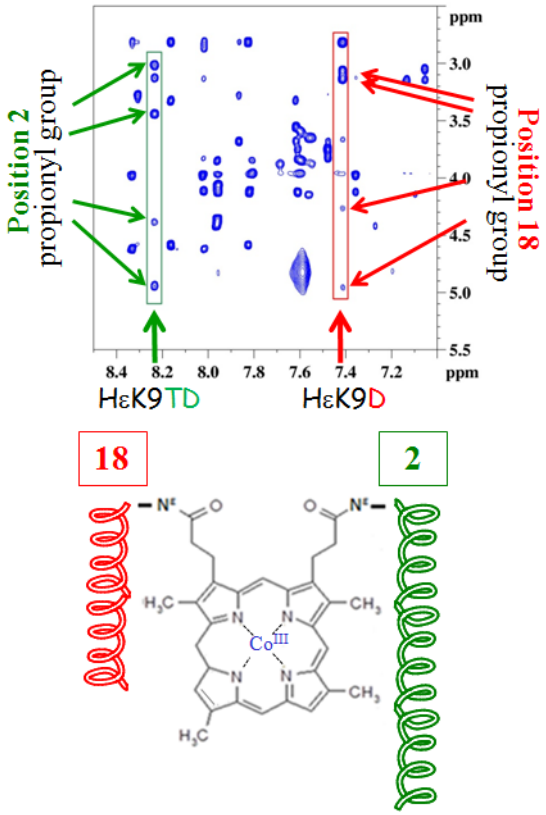


Fig. 35. Details of NMR NOESY (Nuclear Overhauser Effect) spectrum and covalent structure of Co(III) -Mimochrome VI-2U1L peak a.

Peak b

The strategy for assignment was similar to that used for the **peak a**. The elements of secondary structure were first delineated using information provided by CαH conformational shifts and NOE data. All residues, and particularly His[6], showed secondary chemical shifts of αCH protons with negative values, characteristic of a-helical structures. αCH proton chemical shift changes are reported in bar charts in **Figure 36**.

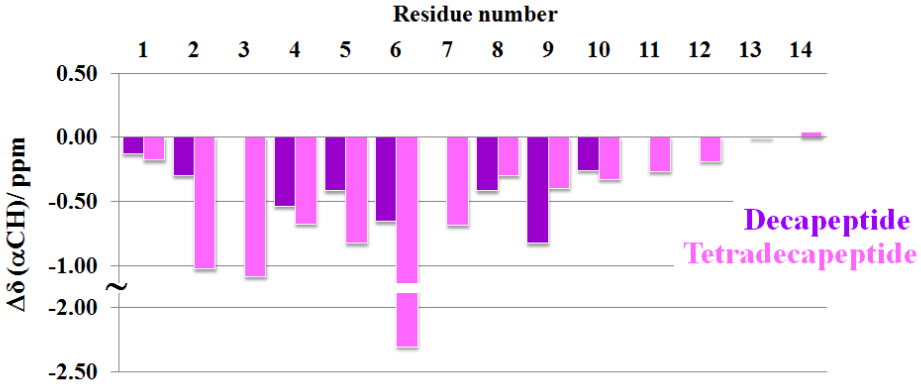


Fig 36. Chemical shift index Δδ(αCH) (δ_{osb} - $\delta_{randomcoil}$) of **peak b**.

This trend was observed for both peptide chains, and it was reasonable to assume that these deviations were due to peptide secondary structure. In particular, it was hypothesized the presence of helical conformation. However, as already reported for **peak a**, we observed a complete pattern of unambiguous NOE connectivities (**see Fig. 37**) for the teradecapeptide chain that allowed us to define, with great confidence, its structure.

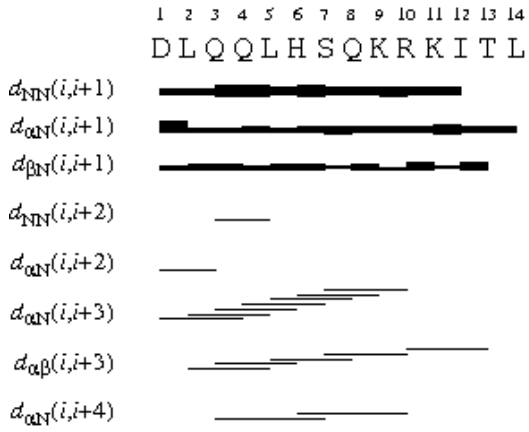


Fig. 37 Co(III)-Mimochrome VI-2U1L **peak** b:summary of NOE contacts. Bar thickness is proportionally related to the NOE intensity.

On the contrary, severe overlaps allowed us to delineate the structure of decapeptide chain only in a qualitative manner. Furthermore, several His^{6} imidazole proton NOEs contacts allowed us to correctly orient the imidazole with respect to the peptide backbone and with the deuteroporphyrin ring. In particular, the histidine exhibits εCH ↔ 8 H, εCH ↔ 10H, εCH ↔ 12 CH_3, δCH ↔ 20H NOE dipolar contacts. These connectivities indicate that the imidazole plane intersects the deuteroporphyrin ring through the 10 and 20 positions of the deuteroporphyrin ring (see **Fig. 32** for deuteroporphyrin numbering).

Finally, the NOE contacts between Lys^{9} sidechain and propionic α,α'- and β,β'-CH_2 protons were crucial to identify the tetradecapeptide chain as the one bound to propionyl group in position 18, and the decapeptide chain the one linked to the propionyl group at position 2 (**Fig. 38)**.

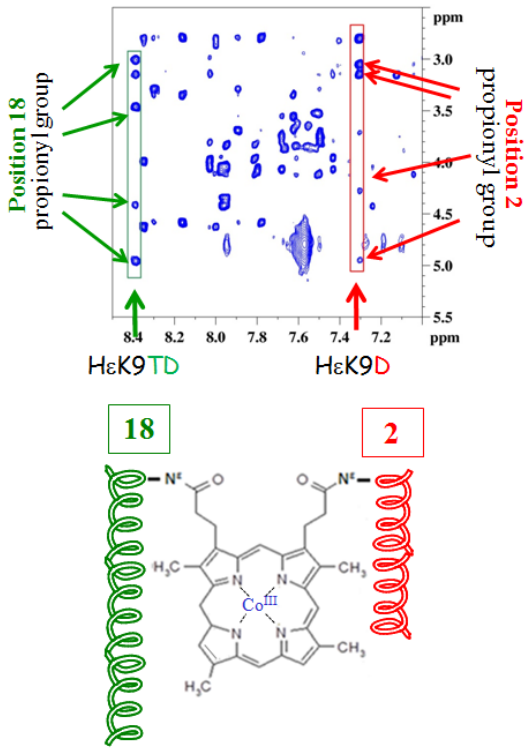


Fig. 38. Details of NMR NOESY (Nuclear Overhauser Effect) spectrum and covalent structure of Co(III)-Mimochrome VI-2U1L peak a.

2.6. Co(III)-Mimochrome VI-2U1L structural calculation

On the basis of the dipolar connectivities, using CYANA software,[135] we were able to determine the solution structures. In particular, for both compound, we were able to solve the structure of the tetradecapeptide-porphyrin moiety, obtaining an ensemble of 20 tightly clustered structures (**Fig. 39**).

The 20 conformers constituting the final family converged to the same topology, with an average total target function of 0.07± 5.16 10^{-2}Å and 0.31 ± 5.16 10^{-2}Å for peak a and b, respectively (table 3 and table 4).

Table 3 NMR statistics of the Co(III)-Mimochrome VI-2U1L **peak a** 20 best NMR structures.	
Rmsd to averaged coordinates Å	
backbone	0.35± 0.15
all heavy	0.93 ±0.19
target function (Å^2)	0.074 ±1.73·10^{-3}
Ramachandran plot statistics (%)	
most favoured regions	91.7
additional allowed regions	8.3
generously allowed region	0.0
disallowed regions	0.0

Table 4 NMR statistics of the Co(III)- Mimochrome VI-2U1L **peak b** 20 best NMR structures.	
Rmsd to averaged coordinates Å	
backbone	0.22± 0.08
all heavy	0.92 ±0.11
target function (Å^2)	0.24 ±3.91·10^{-4}
Ramachandran plot statistics (%)	
most favoured regions	91.7
additional allowed regions	8.3
generously allowed region	0
disallowed regions	0.0

Analysis of the ensemble of the structures by PROCHECK-NMR[136] revealed that 92 % backbone angles fall within the most favored regions of Ramachandran plot, 8% in additional allowed regions , for both peak a and b.

Superposition of the best 20 structures on the mean coordinates gave, for **peak a**, an average root mean square deviation (rmsd) of 0.35±0.15 Å for all backbone atoms and 0.93±0.19Å for all heavy atoms. Similarly the

superposition of the best 20 structures for **peak b** gave, an average root mean square deviation (rmsd) of 0.22±0.08 Å for all backbone atoms and 0.92±0.11 Å for all heavy atoms.

Peak a and **peak b** are characterized by a tetradecapeptide chain with very similar structure. The ϕ and ψ angles from Asp^{1} to Lys^{9} are very close to those expected for a right-handed α-helical conformation (**Fig. 40**).

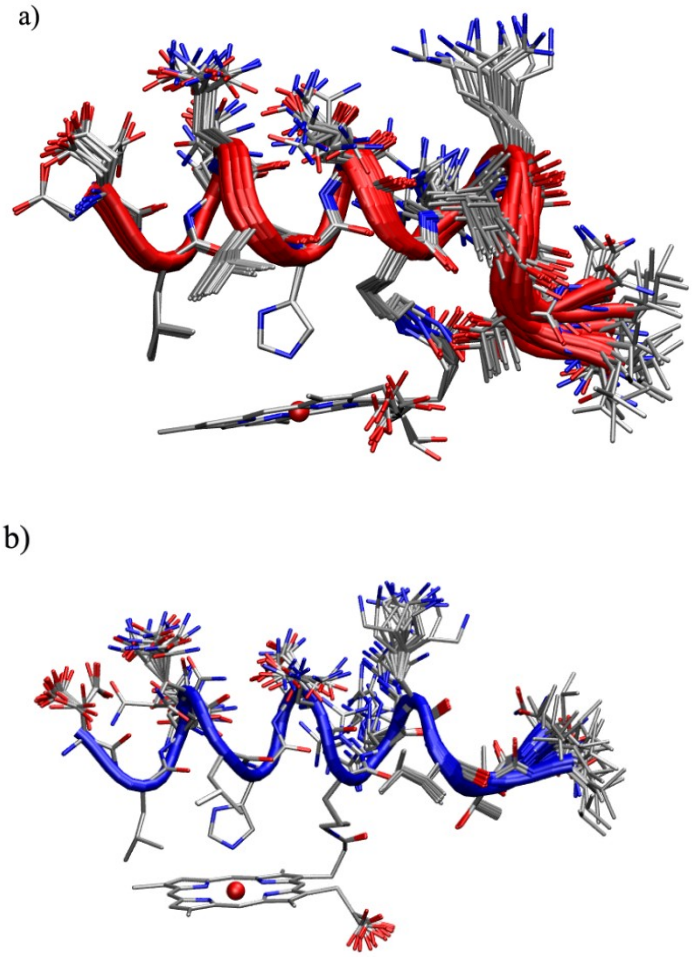


Fig. 39 Solution structure of the tetradecapeptide-porphyrin moiety of Co(III)-Mimochrome VI-2U1L a) **peak a**; b) **peak b**.

The helical structure of the peptide chains is such that Leu^{2}, Leu^{5}, His^{6} and Lys^{9} sidechains are facing the deuteroporphyrin plane. These residues create a partially open hydrophobic cage around the imidazole ring.

In the C-terminal region (segment 11-14), designed to fold in extended conformation, a distortion from α-helical structure is observed. In particular, this distortion is more evident around Arg^{10}-Lys^{11} segment.

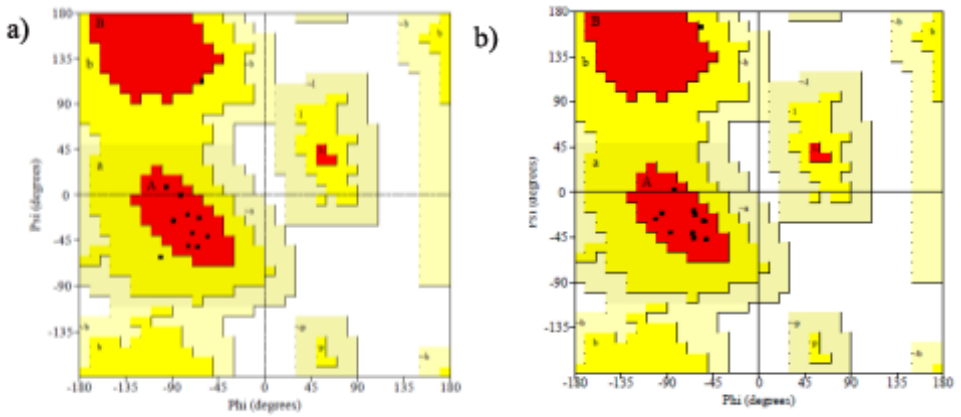


Fig. 40 Ramachandran plot of the tetradecapeptide-porphyrin moiety of Co(III)-Mimochrome VI-2U1L a) peak a; b) peak b.

The dihedral angle values of histidine side chain are χ_1 = -90.9° and -75.5°, χ_2 = 1115.7°and 95°, for **peak a** and **peak b**, respectively. These values were similar to those observed for the side chain conformation adopted by helical histidine-coordinating residue in heme-proteins, and are also in agreement with the hydrogen bond formation between histidine $NH_{\delta 1}$ proton and the backbone carbonyl of the amino acid at position i – 4 (**Fig. 41**).[137]

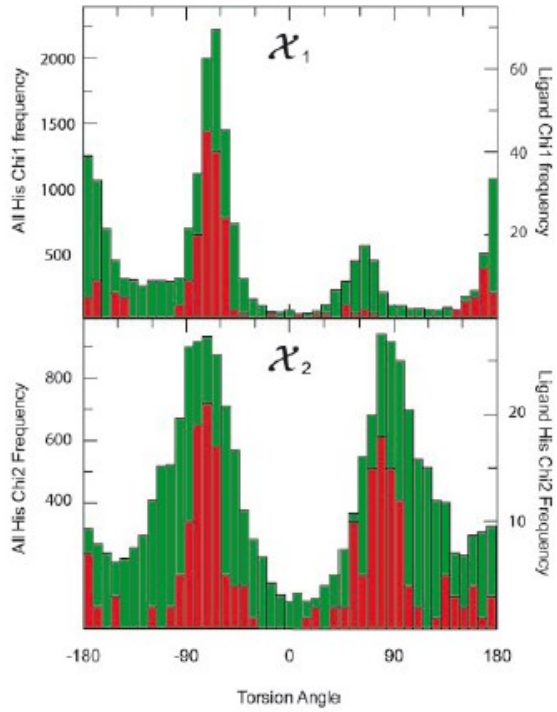


Fig 41. Conformational statistic for histidine derived from the Protein Data Bank select nonredundant protein database. Population distribution derived for all histidines, despicted in green, correspond to the left axis and distribution derived from heme ligand histidines, depicted in red, correspond to the right axis. Copyright © 2004, The National Academy of Sciences.[137]

All peptide bonds are *trans* and all sidechain conformations are staggered. An almost regular CO_i- NH_{i+4} pattern of intramolecular hydrogen bonds is observed for both peptide chains, except for the lack of His^5-CO-Arg^9-NH hydrogen bond and for the presence of two CO_i-HN_{i+3} hydrogen bonds at the N-terminal. In addition, the NH of Lys^9 form hydrogen bonds with the CO groups of His^5. and the NH of Arg^9 and Asn^7, respectively.

3. Development of artificial heme-proteins models with elongated peptide chains: Mimochrome VII

3.1. Design

With the aim of developing catalytically active molecules with a well defined structure, a second strategy was followed in this PhD thesis. This approach involved the design of a new protein framework, characterized by a different organization of the secondary structure elements.

To this aim, the X-ray structure of Co(III)-Mimochrome IV was used as template to engineer a new heme mimetic, Mimochrome VII (**Fig. 42**).

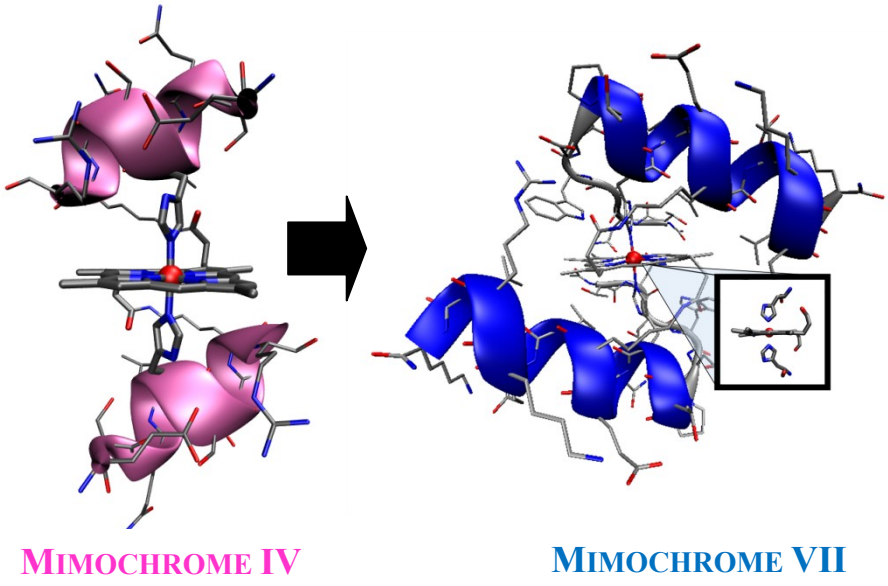


Fig. 42 Molecular structure of Co(III)-Mimochrome IV and molecular model of Co(III)-Mimochrome VII.

Co(III)-Mimochrome IV was fully characterized both in solution and in the solid state. The structural characterization of Co(III)-Mimochrome IV, at the atomic level, confirmed that most of the factors considered in the

design were actually effective in stabilizing the molecules. In the solid state, as well as in solution, Co(III)-Mimochrome IV shows the designed sandwich structure, with the peptide chains in a helical conformation. Indeed, a different orientation of the peptide chains is observed for the solution and X-ray structures. In fact, while the crystallography structure shows two intra-chain ion pair interactions, between the carboxylate side chains of glutamate residues, Glu^{1}, and the guanidine groups of arginine, Arg^{9}, the solution structure shows inter-chain interactions between the carboxylate groups of Glu^{1} on one helix and the guanidine groups of Arg^{9} on the other helix, as intended in the design.

The structural information derived from Mimochrome IV was fruitfully applied for improving the design, and to further stabilize the secondary and tertiary structure. Starting from the Co(III)-Mimochrome IV X-ray structure, we elongated the peptide chain and modified the sequence by introducing several residues able to make inter-chain and intra-chain interactions. The design process lead to the development of a new heme-protein mimetic, Mimochrome VII, with the following features (**Fig. 42**):

- it is made up of two equal 17-residue peptide chains anchored to the heme group. Each chain bears an histidine residue that acts as axial ligand to the heme central ion, providing a symmetrical six-coordinate coordinated metal binding site;
- the chains fold into two secondary structure motifs (**Fig. 43**):
  - the C-terminal residues (from 6 to 17) were modeled to stabilize an α-helix;
  - residues at N-terminus (from 1 to 5) were modeled in an extended conformation, folding back to interact with the helical part. This segment will be useful in the design of

asymmetrical penta-coordinated models, because it may ensure further protection of the heme from bleaching, during the catalytic cycle.

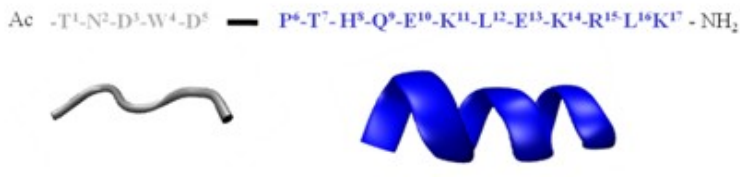


Fig. 43: schematic representation of Mimochrome VII secondary structure.

To induce the desired secondary structure, amino acid with high helix propensity were inserted at the C-termini, amino acid with low helix propensity were placed at N-termini (**Fig 44**);

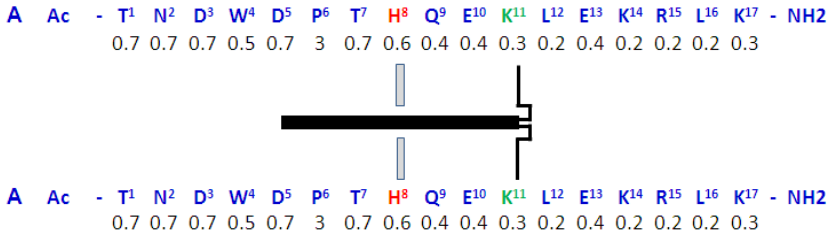


Fig. 44 Schematic representation of Mimochrome VII molecule. The metalloporphyrin moiety is depicted as a black line. Below each amino acid, helix propensity according to Pace and Scholtz is reported.[121] The results are presented as Δ(ΔG) values relative to Ala, which has been set to zero because it is usually the amino acid with the most favorable helix propensity. The coordinating histidine residues are highlighted in red, the lysine residues covalently linked to DP IX are highlighted in green.

The peptide chains were designed to both fold in the intended tertiary structure and create the desired coordination shell around the heme. The sequence was modified in order to introduce several residues able to make inter-chain and intra-chain interactions in strategic positions. Computational methods based on molecular mechanics were used in order to refine the structure into a local energy minimum, and through an iterative process of

design and energy minimization, the design process culminates in the development of a new heme-protein mimetic, with the following features:

1. the amino acid composition was designed to contain hydrophilic residues in all the solvent exposed positions (Glu^{10}, Glu^{13}, Lys^{14}, Lys^{17}) and hydrophobic residues were placed in the positions facing the porphyrin moiety (Leu^{12}, Leu^{16}).

2. Glu and Lys residues were placed at positions 10, 13 and 14, 17 of the sequence, respectively, in order to engineer two helix-stabilizing i, i+4 ion pair interactions in each helix.

3. a tryptophan residue was placed at the N-termini extended segment of each chain. By forming hydrophobic interaction with the porphryn, these tryptophan residues favour the folding back of the peptide chains.

4. two inter-chain ion pair were introduced (**Fig. 45**), to avoid the formation of multiple diastereomeric forms, which were observed in early Mimochrome mimetics [119] In fact, due to the linker flexibility between the peptide and the porphyrin cofactor, two different orientations of the peptide chains are allowed: each peptide chain can be arranged above or below the porphyrin plane, giving rise to different diasteromeric configurations around the iron center. To favor the formation of one of these configurations,[124] two ion pair interations have been introduced between arginine and aspartic acid residues, that were placed

according to intrinsic helix dipole, with the negative residue at the N-terminus of the helix (Asp^5), and the positive residue at the C-terminus (Arg^{15});

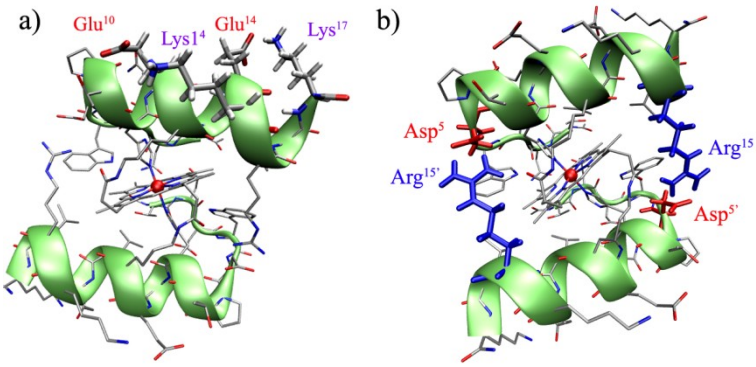


Fig. 45 Fe(III)-Mimochrome VII model. Intra-chain a) and inter-chain b) ion pair interactions are highlighted.

5. N- and C-terminal ends were acetylated and amidated, respectively, to avoid the presence of end charges that might affect the helix stability.

3.2. Synthesis and purification

The new compound was synthesized similarly to other Mimochromes, by using previously developed protocols. The synthetic strategy is schematically reported in **Fig. 46**. The 17 amino acid peptide was synthesized using solid-phase methods with 9-Fluorenylmethoxycarbonyl (Fmoc) protocol. After selective deprotection of Lys^{11}, the peptide was cleaved from the resin, without removing the other side chain protecting groups. It was then coupled to deuteroporphyrin IX in solution. The resulting miomicohrome VII apo form was completely deprotected and purified by

RP-HPLC. Finally, iron and cobalt ions were inserted into the porphyrin ring, according to literature procedures. Electrospray ionization mass Spectrometry (ESI-MS) and ion trap (IT) - time-of-flight (TOF) mass spectrometry (LCMS-IT-TOF) were used to confirm the identity of the products. Experimental details on synthesis, RP-HPLC purification and ESI-MS spectra are reported in experimental section.

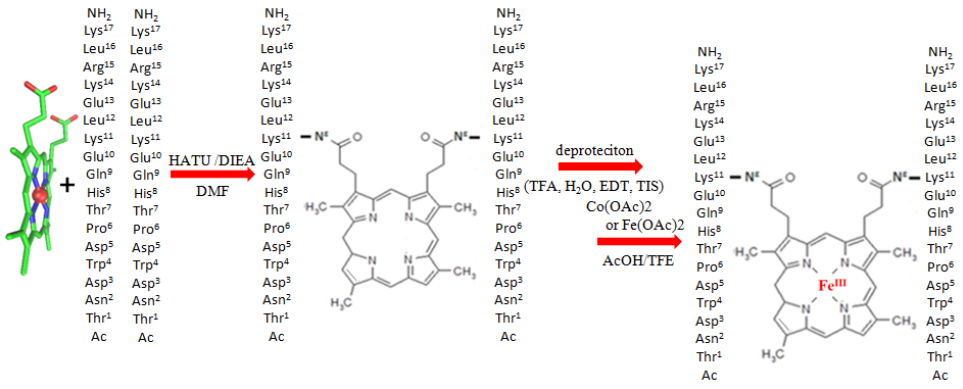


Fig. 46 Synthetic strategy.

3.3. CD Spectroscopy

The conformational features of Co(III)-Mimochrome VII and Fe(III)-Mimochrome VII complexes were investigated using CD spectroscopy. Fig. 44 shows the UV-region CD spectra of both complexes in 10.0 mM phosphate buffer solution (pH=7), and in phosphate buffer solution (pH=7) containing 30% TFE (v/v) (**Fig. 47**). Table 5 reports their CD parameters along with the CD parameters of Fe(III)- and Co(III)-Mimochrome IV complexes.[123]

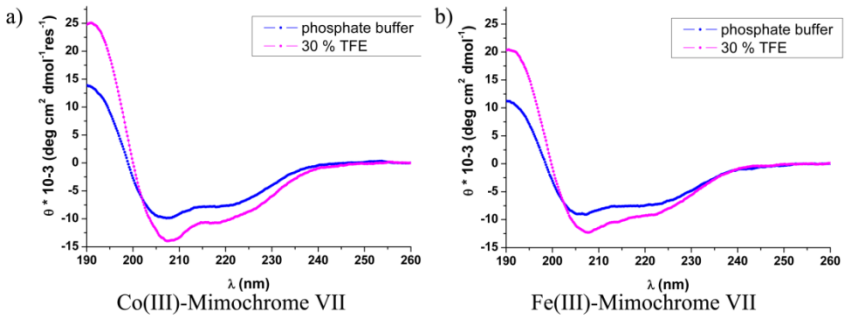


Fig. 47: CD spectra in the far/UV region of a) Co(III)-Mimochrome VII in phosphate buffer, 10 mM, pH 7(blu line) and in phosphate buffer, 10 mM, pH 7/30% TFE (v/v) (pink line); b) Fe(III)-Mimochrome VII in phosphate buffer, 10 mM, pH 7(blu line) and in phosphate buffer, 10 mM, pH 7/30% TFE v/v) (pink line).

The CD spectral behaviors for both Co(III)- and Fe(III)-Mimochrome VII complexes clearly indicate the peptides to be in helical conformation, as designed. The shapes and relative intensities of the minima around 222 and 206 nm, as well as the λ_0 crossover and maximum around 190 nm, indicate the peptides to be helical in both Fe(III)- and Co(III)- complexes. The helical content slightly increases upon TFE addition (up to 30%) and it is quite comparable to that of the corresponding Mimochrome IV complexes (see Table 5). This finding clearly indicates that the amino acid substitutions made in the Mimochrome IV sequence did not perturb the helical

conformation of the peptide chains, but they play a favorable contribution in the helix formation.

Table 5 CD parameters of Mimochrome VII compared with the values for Mimochrome IV.

	$[\theta]_{min}$ *10^{-3} (λ, nm)	λ_0 (nm)	$[\theta]_{222}$ *10^{-3} (λ, nm)	$[\theta]_{max}$*10^{-4} (λ, nm)
Fe(III)- Mimochrome VII				
10 mM phosphate buffer pH 7	-9.1 (207)	199	-7.3	1.1 (190)
phosphate buffer /TFE 70/30 (v/v)	-12.3 (208)	200	-9.2	2.0 (190)
Co(III)-Mimochrome VII				
10 mM phosphate buffer pH 7	-9.8 (208)	199	-7.4	1.4 (190)
phosphate buffer/TFE 70/30 (v/v)	- 14.0(208)	200	-9.7	2.5 (191)
Fe(III)- Mimochrome IV				
phosphate buffer/TFE 70/30 (v/v)	-11.2 (207)	200	-9.5	1.7 (193)
Co(III)- Mimochrome IV				
phosphate buffer/TFE 70/30 (v/v)	-15.2 (207)	200	-11.4	2.2 (192)

The θ_{222} values for both Co(III) and Fe(III) complexes are lower than that expected for an almost complete helical conformation. Even though the θ_{222} value can be correctly applied to calculate the helix percentage in proteins, it fails to estimate the helix contents in small peptides,[138] for which the θ ratio of the minimum at 222 nm to the minimum at shorter wavelengths, the position of this last minimum and the crossover wavelength λ_0 are all indicative of helix propensity. Further, the interactions between the heme transitions and those of the peptide backbone amide dipoles may influence the far-UV CD spectra. The value of [θ]ratio, the [θ] ratio of the minimum at 222 nm to the minimum at shorter wavelengths, is around 0.8. Typically, in the a-helix the [θ] ratio value is around unity. The observed values of the [θ]ratio could be interpreted as a small 3_{10}-helical distortion, as reported in literature.[139]

3.4. UV-vis spectroscopy

UV/Vis spectroscopy was used to determine the coordination geometry, the spin and oxidation states of the metal ion inserted into the porphyrin ring. **Figure 48** shows the UV/Vis spectra of Fe(III)-Mimochrome VII (**a** and **b**) and Co(III)-Mimochrome VII (**b** and **c),** in 10.0 mM phosphate buffer solution, pH 7, and 10 mM phosphate buffer, pH 7/TFE 70/30 (v/v), respectively.

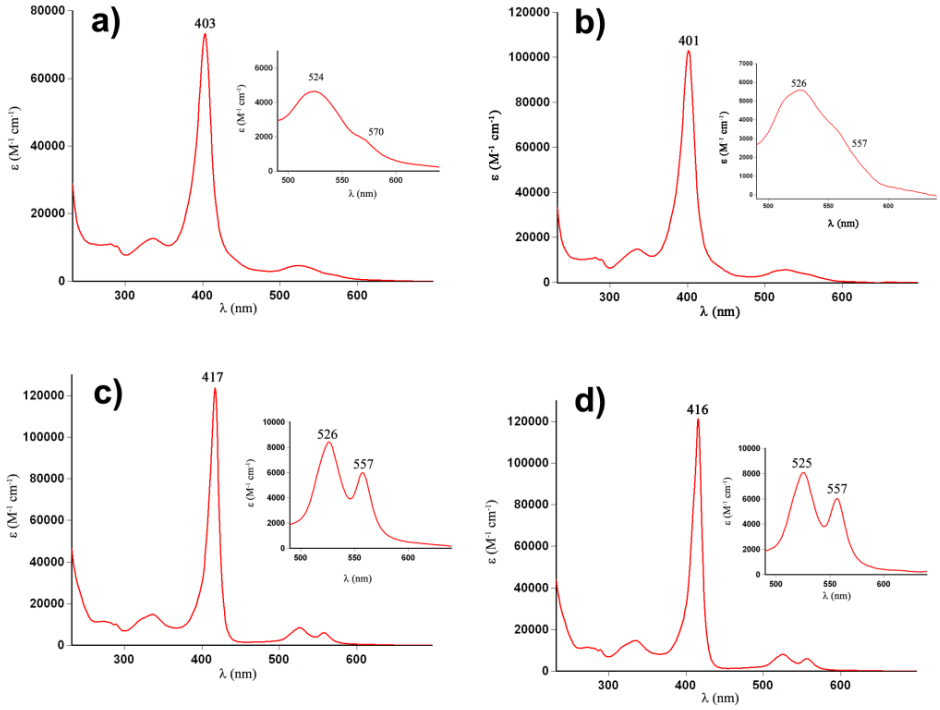


Fig. 48 UV-vis spectra of Fe(III)-Mimochrome VII a) in 10 mM phosphate buffer, pH 7, b) 10 mM phosphate buffer, pH 7/ TFE 70/30 (v/v); Co(III)-Mimochrome VII in 10 mM phosphate buffer, pH 7, d) 10 mM phosphate buffer, pH 7/ TFE 70/30 (v/v).

The position and the relative intensities of the Soret, β and α bands give useful information on the metal ion coordination state.[6,45] For the iron derivative at pH 7, the positions of the Soret and low energy bands at 403 (ε=7.4 × 10^4 $M^{-1}cm^{-1}$) and 524 nm (ε=4.6 × 10^3 $M^{-1}cm^{-1}$), respectively, with

a shoulder at 570 nm are indicative of a ferric low-spin state, with a bis-His axial coordination.[6,45] For the cobalt complex at pH 7, the positions of the Soret, β and α bands at 417 nm (ε=1.2 10^5 × $M^{-1}cm^{-1}$), 526 nm (ε=8.4 10^3 × $M^{-1}cm^{-1}$) and 557 nm (ε=6.0 10^3 × $M^{-1}cm^{-1}$) respectively, and the extinction coefficient ratio between the β and α bands are characteristic of an octahedral Co(III).[119,122,140,141]

All the spectral data (see table 6) are very similar to those found for the Fe(III) and Co(III) complexes of Mimochrome IV.

Table 6. UV-vis spectral data of Fe(III)-Mimochrome VII and Co(III)-Mimochrome VII

Specie	pH	λ_{Soret}, nm (ε, M^{-1} cm^{-1})	$\lambda_{Visible}$, nm) (ε, M^{-1} cm^{-1})
Fe(III)-Mimochrome VII			
10 mM phosphate buffer	7	403 ($7.4*10^4$)	524, 570 ($4.6*10^3$, $1.8*10^3$)
phosphate buffer /TFE 70/30 (v/v)	7	401 ($1.0*10^5$)	526, 557 ($5.6*10^3$, $3.5*10^3$)
Co(III)-Mimochrome VII			
10 mM phosphate buffer	7	417 ($1.2*10^5$)	526, 557 ($8.4*10^3$, $6.0*10^3$)
phosphate buffer /TFE 70/30 (v/v)	7	416 ($1.2*10^5$)	525, 557 ($8.1*10^3$, $6.0*10^3$)
Fe(III)-Mimochrome IV			
phosphate buffer	7	401 (1.0*105)	522, 560 (6.2*103)
Co(III)-Mimochrome IV			
phosphate buffer	7	415 (1.6*105)	525, 556 ($1.2*10^4$, $8.2*10^3$)

In order to determine the coordination properties of the Fe(III) complex, the UV-vis spectra at different pH (from pH 1 to 7) were recorded in water/TFE solution 70/30 (v/v). **Fig.49** shows the UV/Vis spectra of Fe(III)-Mimochrome VII at pH 7 (blue line), and at pH 1.4 (red line). Figure 47 reports the plot of the molar extinction coefficient intensities of the band at 388 nm (red line) and 401 nm (blue line) as a function of the pH.

At pH 1.4, the electronic spectrum shows the Soret band at 388 nm (ε ~ 1.0 × 10^5 M^{-1} cm^{-1}), the Q bands at 501 and 528 nm (ε ~ 6.4 × 10^3 M^{-1} cm^{-1} and ~ 5.7 × 10^3 M^{-1} cm^{-1}), and a CT (charge-transfer) band at 617 nm (ε ~

2.3×10^3 M^{-1} cm^{-1}). This spectrum is in indicative of a high-spin ferric state, with two weak ligands (i.e., water molecules) axially coordinated to the iron

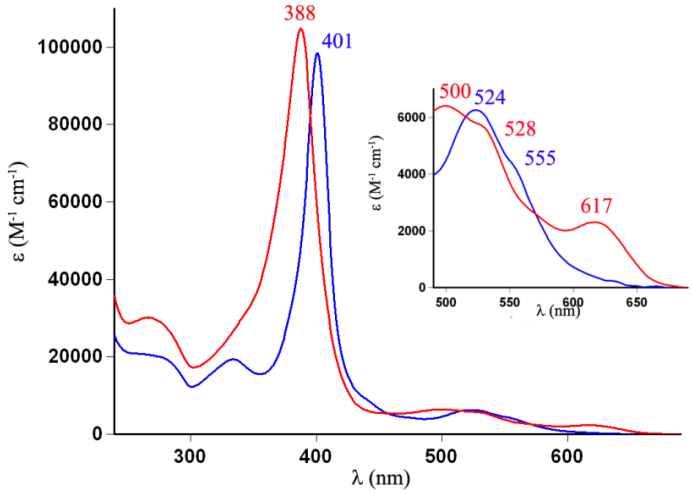


Fig. 49 Fe(III)-Mimochrome VII Uv-vis spectra at pH 7 (in blue) and at pH 1.4 (in red).

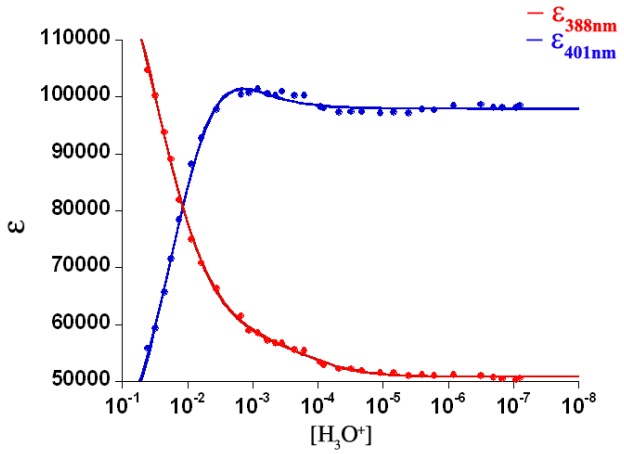


Fig. 50 The dependence of Fe(III)-Mimochrome VII molar extinction coefficient (ε) at 388 nm (in red) and at 401 nm (in blue) on pH. The fitting of the experimental data yielded the following values: pKa1 =1.6 and pKa2 =3.9.

By raising the pH above 4.5, the Soret band maximum shifts at 401 nm ($\varepsilon \sim 9.8*10^4$), the Q band move at 524 and 557 nm ($\varepsilon \sim 6.2 *10^3$ and $4.2 *10^3$, respectively) and CT band at 618 nm disappears. This observation is

indicative of transition from a predominantly high-spin ferric state to a predominantly low-spin ferric state, typical of a bis-His coordination (characterized by a strong ligand field on the iron porphyrin).

Two midpoints are observed at pH 1.5 and 3.9. In the pH range 1.4-2.5 (see **Fig. 50)** the presence of an isosbestic point (**Fig. 51**) suggests that only two species are predominantly involved in the equilibrium. In analogy with previously characterized Mimochrome analogues, we suggest the first pKa (pK_a = 1.50) corresponds to the apparent protonation of the axially coordinated histidines (pK_{app}), as defined by Kennedy.[88]

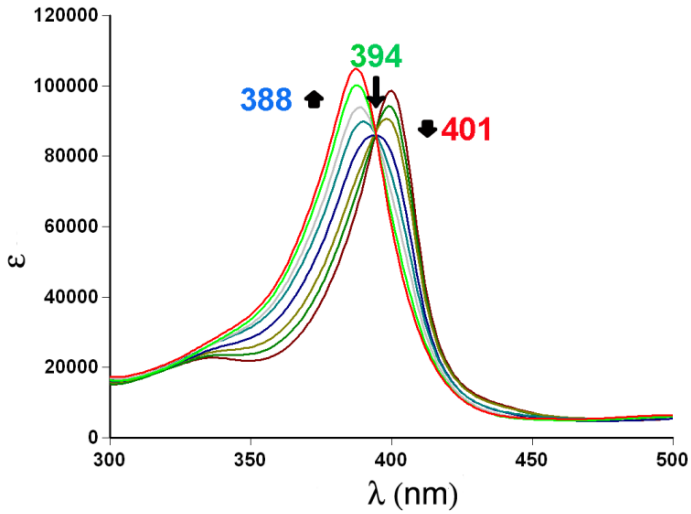


Fig. 51 UV-vis spectra of Fe(III)-Mimochrome VII in the pH range 1.4 - 2.5. Arrows indicate changes at 401 and 388 nm from alkaline to acidic pH. An isosbestic point appears at 394 nm.

The very low value of the His pKa, point out very strong coordination of the axial histidine residues to the iron, and may indicate a highly hydrophobic local environment. This is in agreement with the designed model, which shows two leucine residues (Leu^{12} on each chain) that forms

hydrophobic interactions with the porphyrin, surrounding the histidine residues and creating a highly hydrophobic local environment.

The pK_{a2} of Fe(III)-Mimochrome VII is more difficult to interpret. Since no major change in spin state occurs during this equilibrium, we suggest this pK_{a2} does not correspond to changes in the first coordination shell of the heme iron. More reasonably, the observed modification of the electronic spectrum are related to protonation equilibrium involving amino acid in the second coordination shell (see **Fig. 52**). In particular, an aspartic residue (Asp^3), that was predicted to form hydrogen bond with the coordinating histidine, is suggested to be responsible for this transition (3.65 is the pKa for Aspartic acid side chain in free amino acid).[142]

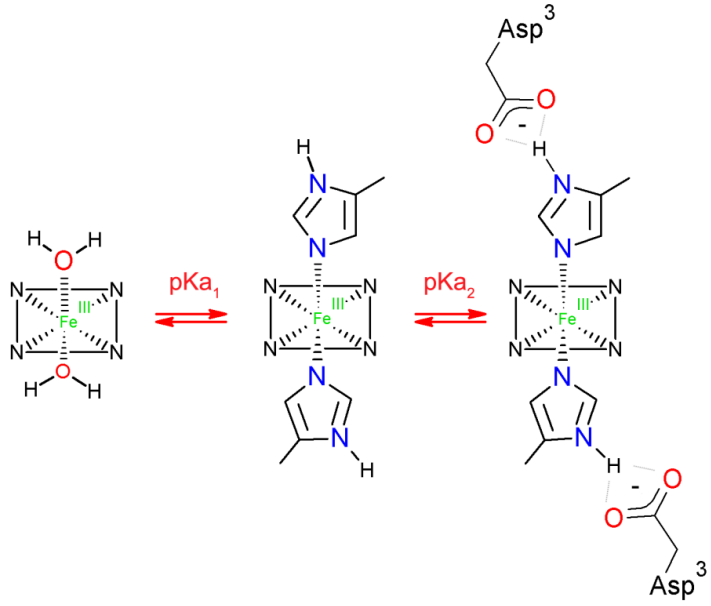


Fig. 52 Proposed equilibria and species involved in the pH tritation of Fe(III)-Mimochrome VII.

3.5. Co(III)-Mimochrome VII: NMR spectroscopy

The diamagnetic cobalt complexes of Mimochrome VII was analyzed by Nuclear Magnetic Resonance. The NMR analysis was carried out in H_2O/TFE 50/50 (v/v), **Fig. 53** shows the ^{1}HNMR spectrum of Co(III)-Mimochrome VII complex.

There are two set of resonances for the deuteroporphyrin protons in ≈1:3 ratio, suggesting the presence of two species in solution. We named **compound 1 (c.1**) the most abundant species, and **compound 2 (c.2)** the other one.

Due to resonance overlaps, we were not able to solve the three dimensional structure of **compound 2, c.2** and the complete structural characterization was performed only on c**ompound 1, c.1**.

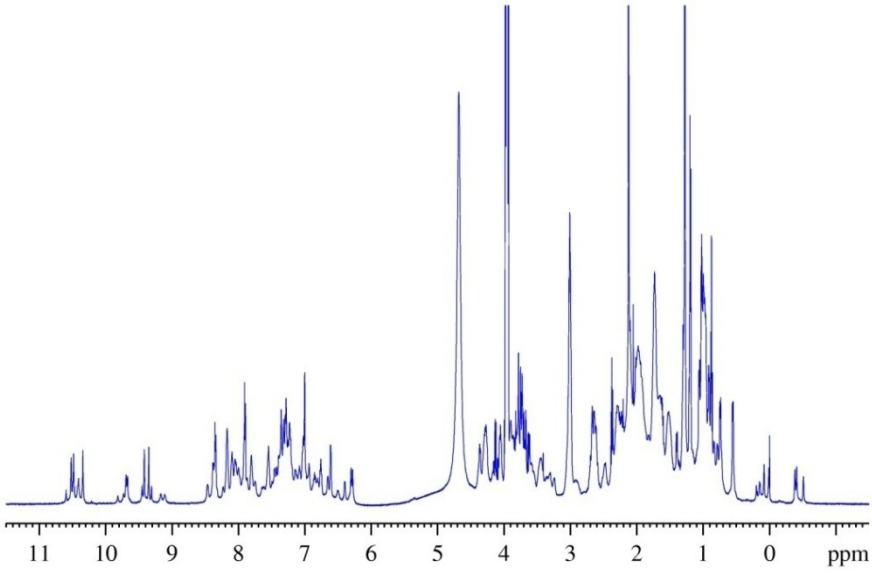


Fig. 53. Co(III)-Mimochrome VII 1D ^{1}H spectrum.

3.6. Co(III)-Mimochrome VII c.1: resonances assignment and analysis

Resonances assignment was accomplished by 2D experiments (NOESY, TOCSY, DQF-COSY) using standard methodologies. All the peptide and deuteroporphyrin proton resonances were unambiguously assigned. The assignment of deuteroporphyrin resonances was easily performed following the NOE connectivities around the macrocycle: the 5H proton was directly identified, as it is the only it is the only porphyrin proton in dipolar contact with two methyl groups (3-CH3 and 7-CH). Thanks to dipolar connectivity with methyl 7-CH3 it was easy to identify the pyrrole 8H resonance. This lead to immediate identification of the β-meso proton at position 10H, and, for exclusion, of pyrrole proton 13H and of the β-meso proton 5H. The β-meso proton 20H was finally identified, as it shows strong dipolar connectivities with the propionic α,α'-CH_2 and β,β'-CH_2. In particular, it was possible to distinguish the resonances of the αα'CH2 and β,β'CH_2 protons of the propionic group at position 2 of the deuteroporphyrin ring from those of the propionic at position 18 (see **Fig. 32** for deuteroporphyrin numbering). Dipolar contacts were observed between these protons and the amino acidic residues, providing unambiguous assignment of the two peptide chains.

We observed two sets of resonances for the peptidic part, with severe overlap for some of the homologous residues belonging to the two different peptide chains. These overlaps are more marked in the N-terminal region (Thr^1, Asn^2, Asp^3, Trp^4).

Lysine 9 is covalently bound to the deuteroporphyrin propionyl groups. As already observed for other Mimochrome analogues, side chain

protons of this residue are upfiled shifted respect their random coil values, due to the ring current effects.

Histidine side chain protons were unequivocally assigned. In the TOCSY experiment, the δ-CH signals at δ = -0.38 and -0.40 ppm are correlated to ε-CH imidazole protons at δ = 0.02 and 0.09 ppm, respectively. These last resonances are in strong dipolar contact with resonance at δ = 9.17 and 9.11 ppm, which were assigned to the imidazole δ-NH of the two histidines. The imidazole Nε atom is unprotonated and axially coordinate the Co(III) ion.

Imidazole δ-CH and ε-CH protons have different connectivities with the deuteroporphyrin protons, thus indicating that the two histidines are characterized by different orientation with respect to the porphyrin plane. One imidazole ring presumably lies in a plane that is orthogonal to the deuteroporphyrin plane and oriented almost parallel to the N_{21}-N_{23} direction, while the other imidazole plane is almost parallel to the N_{22}-N_{24} direction (**Fig. 54**).

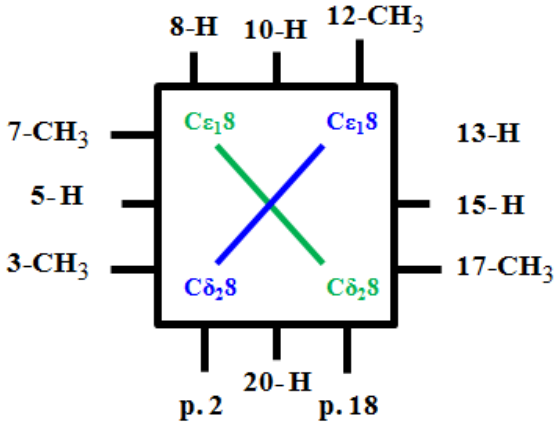


Fig. 54 Histidine axial ligands orientation in Co(III)-Mimochrome VII compound 1.

Table 7 Dipolar connectivities between histidine imidazole protons and deuteroporphyrin IX protons			
Histidine (p.2)	**DPI X**	**Histidine (p.18)**	**DPI X**
Hδ2	20 H	Hδ2	20 H
	5H		15 H
	3-CH3		13 H
	α/α'CH2		17- CH3
	β/β'-CH2		α/α'CH2
			β/β'-CH2
Hε1	10 H	Hε1	10 H
	15 H		8 H
	13 H		7- CH3

Both peptide chains showed a very similar pattern of NOE interactions, which strongly support the presence of a well organized structure. A right-handed helical conformation could be inferred from Pro[6] to Leu[16] on the basis of the characteristic short-range NH_i-NH_{i+1}, αCH_i-NH_{i+1}, βCH_i-NH_{i+1}, and medium-range αCH_i-βCH_{i+3}, αCH_i-NH_{i+3}, αCH_i-NH_{i+4}. A non-regular structure occurs at N-terminal end (residues 1-5), which displays only few of the connectivities expected for a helical conformation. The αCH chemical shifts were additional indicators of helical structure (Fig. 51). All residues and particularly His[8], showed αCH resonances significantly up-field shifted, relative to their random coil values (**Fig. 55**). Since this trend was observed also for the N-terminal residues, it is reasonable to assume that these deviations are due not only to peptide secondary structure but also to the deuteroporphyrin ring current effect.

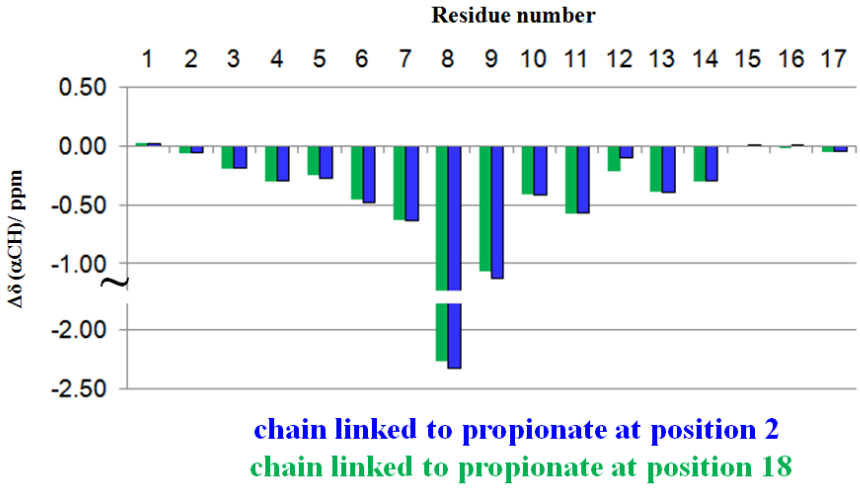


Fig. 55 chemical shift index Δδ(α-CH) (δ_{osb} - $\delta_{randomcoil}$) of compound 1.

3.7. Co(III)-Mimochrome VII c. 1:solution structure from NMR data

The solution structure of Co(III)-Mimochrome VII **c.1** was solved by using CYANA software,[135] on the base of 554 dipolar connectivities (**Fig. 56**). Typical CYANA runs were performed and the 20 best structures are tightly clustered.

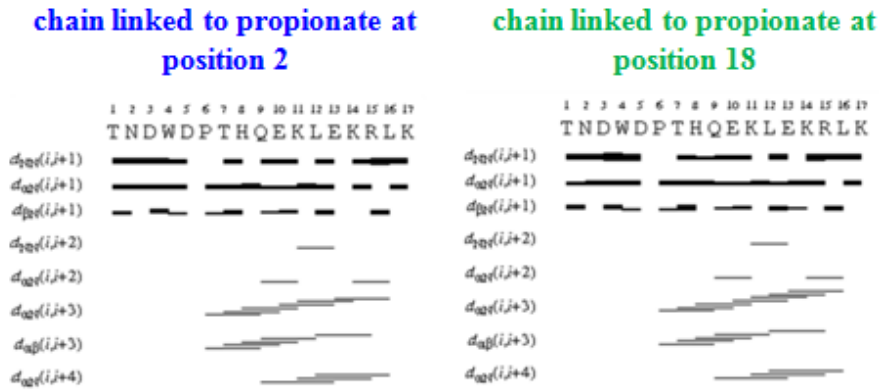


Fig. 56 summary of NOE contacts typical of secondary structure in Co(III)- Mimochrome VII spectrum. Each contact is represented by an horizontal line which connects the involved residues. The lines thickness represents the contacts intensity.

Analysis of the ensemble of the structures by PROCHECK-NMR[136] revealed that 75.2% of backbone angles fall within the most favored regions of Ramachandran plot, 23.2% in additional allowed regions and 1.6% in the generously allowed regions.

Superposition of the best 20 structures on the mean coordinates gave, for Co(III)-Mimochrome VII **c.1**, an average root mean square deviation (rmsd) of 0.77 ±0.15 Å for all backbone atoms and 1.19 ±0.12 Å for all heavy atoms.

Table 8 NMR statistics of the Co(III)-Mimochrome VII 20 best NMR structures.

Rmsd to averaged coordinates Å	
backbone	0.77 ± 0.15
all heavy	1.19 ± 0.12
target function (Å^2)	0.65 ±5·10^{-2}
Ramachandran plot (%)	
most favored regions	74.2
additionally allowed regions	22.6
generously allowed regions	3.2
disallowed regions	0.0

The two peptide chains in Co(III)-Mimochrome VII adopt right-handed helical conformation from Pro^{6} to Leu^{16} (see **Fig. 41** for residue numbering), while a less regular conformation is observed in the N-terminal region (**Fig. 57**). As designed, the global structure is stabilized by the presence of interactions between hydrophobic side chains and the deuteroporphyrin ring. In particular, Leu^{12} and Trp^{4} residues are facing the deuteroporphyrin plane, and they form hydrophobic interactions with the macrocycle. The side chains of hydrophilic residues (Glu^{10}, Glu^{13}, Lys^{14}, Lys^{17}) are solvent exposed.

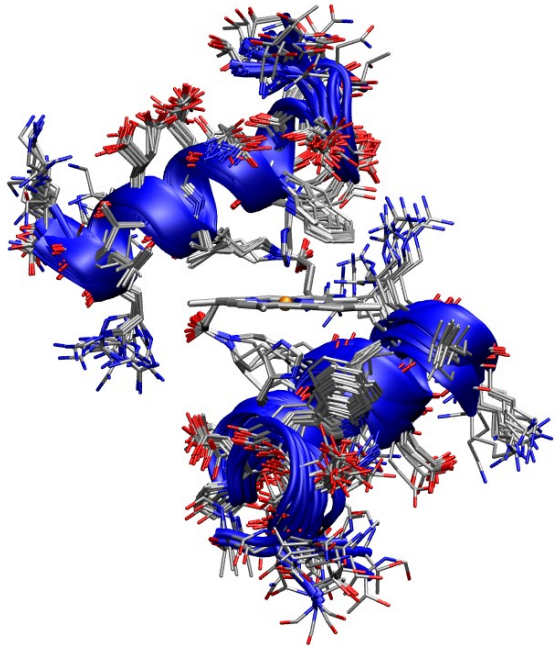


Fig. 57 Co(III)-Mimochrome VII three dimensional NMR structure, represented as a bundle of twenty structure.

3.8. Co(III)-Mimochrome VII compound 2: preliminary resonances analysis

We also analyzed Co(III)-Mimochrome VII **c.2**. Unfortunately, due to severe resonance overlaps with compounds 1, we were not able to solve the NMR structure of this species.

However, the qualitatively analysis of dipolar contacts together with the αCH chemical shift (up-field shifted) suggested the presence of a helical conformation in the 6-17 region (see **Fig. 58**).

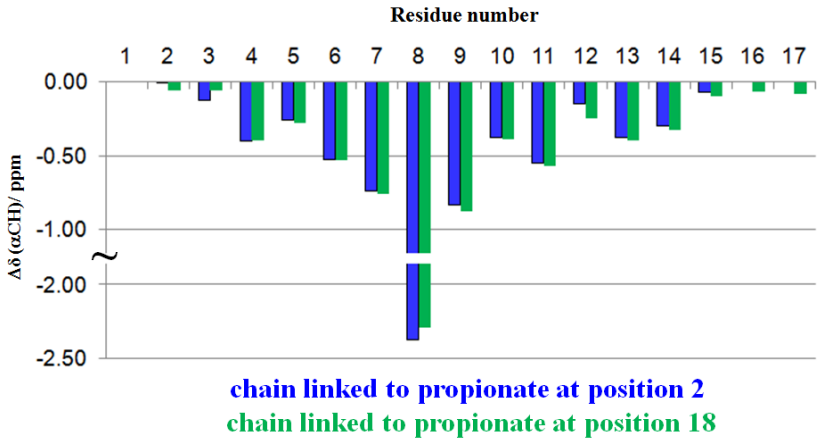


Fig. 58 chemical shift index Δδ(α-CH) (δ_{osb} - $\delta_{randomcoil}$) of compound 2

The most noticeable difference between the two compounds is the orientation of the histidine rings with respect to the heme plane, as indicated by the different connectivities exhibited by imidazole δ-CH and ε-CH protons with respect to the deuteroporphyrin protons. A comparision of the imidazole orientation o in compound 1 and compound 2, as deduced from NOE connectivities, is schematically reported in **Fig. 59**.

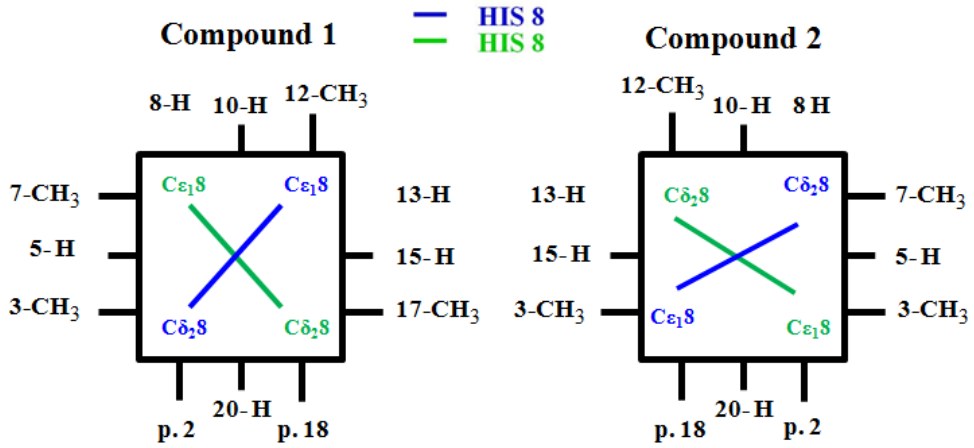


Fig. 59: different orientation of imidazole rings in compound 1 and compound 2.

Different orientation of the imidazole rings with respect to the porphyrin are indicative of a different topological form, since the orientation of the histidine side chain is a key factor in determining the orientation of the peptide chain with respect to the porphyrin plane.

3.9. Design of a new model with improved peroxidase like activity: Mimochrome VII asym.

The structural information obtained on Co(III)-Mimochrome VII were very useful in order to re-design a new model with peroxidase-like activity, which couple a well defined three dimensional structure with high catalytic performances. Therefore, the NMR structure of Mimochrome VII was used as a template to design a new complexes with an asymmetrical active sites (**Fig. 60**).

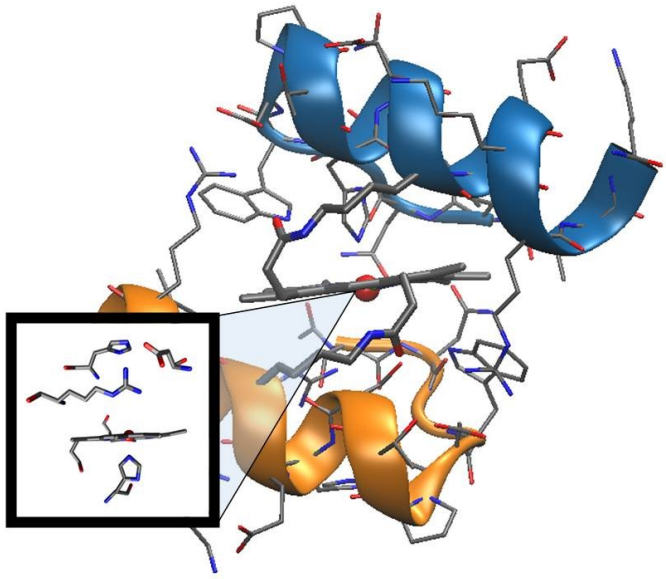


Fig. 60 Molecular model of Fe(III)-Mimochrome VII asym.

This new model is characterized by the engineering into Mimochrome VII of a distal cavity, mimicking the HRP distal pocket. This cavity contains an arginine residue (Arg^{12}), mimicking Arg^{38} of Horseradish Peroxidase (HRP), and a non-coordinating His residue (His^{9}) to mimic His^{42} of HRP (**Fig. 61**).

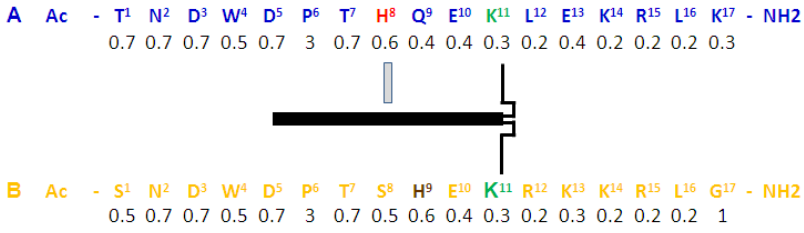


Fig. 61 Fe(III)-Mimochrome VII asym. sequence. The metalloporphyrin moiety is depicted as a black line. Below each amino acid, helix propensity according to Pace and Scholtz is reported.[121]

4. Artificial heme-proteins: determination of axial ligand orientations through paramagnetic NMR shifts

To date, several heme protein models in the Mimochromes family have been developed, and some of these models have been extensively characterized. However, so far, structural data on Mimochrome active site geometry were obtained through the characterization of the diamagnetic Co(III) complexes. A paramagnetic ion modifies the properties of the nuclear spins coupled to it, and dramatically alters the NMR spectra. The coupling can occur through both chemical bonds and space, affects chemical shifts and relaxation rates. Specifically, the relaxation rates increase due to coupling with unpaired electron(s), and this phenomenon induces, in a distance-dependent manner, line broadening of vicinal nuclear spins, even beyond signal detection.[143,144] Nevertheless, the paramagnetic effects are an important source of structural information, as they provide distance and geometric details about the environmentof the paramagnetic center in metalloproteins.

In order to get information about the orientation of the axial ligands, the NMR characterization of Mimochrome IV iron(III) paramagnetic complex was performed. In particular, by combining an empirical equation, proposed for bis-histidine ferriheme-proteins, with experimental data obtained for Fe(III)–mimochrome IV, we obtained information about histidine orientations.[145,146]

In its essence, this methodology take advantage of the understanding of the relationship between the electronic structure of the heme, inferred by the paramagnetic shifts of its nuclei, and the orientation of the axial ligand(s).

The heme of the low-spin hemeproteins has been characterized extensively, and semi-empirical and empirical equations have been developed to correlate

the hyperfine shifts of the heme-methyl resonances with the orientation of the iron axial ligand(s).[147-153]

With the aim to extend these methods to peptide-based artificial metalloproteins, and to fully characterize the active site structures of mimochromes, during this PhD research project, in collaboration with the Inorganic Biochemistry and NMR group of the Instituto de Tecnologia Química e Biológica of Lisbona, the empirical equation proposed by Turner was applied to Fe(III)-Mimochrome IV complex.[145]

Within the family of Mimochromes, Mimochrome IV is a simple, structurally defined heme-protein model, with two histidine residues which act as axial ligands to the metal ion, that was completely characterized. In particular, the three-dimensional structure of its Co(III) complex was determined by NMR spectroscopy and X-ray diffraction. The solution and solid state structures share numerous features and conform well to the design, adopting the designed sandwich structure, with the Δ configuration of the Co(III) ion, and have been used as framework for the engineering of new models with improved properties, in this and other works. Due to its simple and well defined structure, Mimochrome IV molecule can be a useful tool for exploring the delicate mechanisms that control the heme functions, such as redox potential. Indeed, the heme redox potential of Mimochrome IV falls within the range observed for natural cytochromes, which present a bis-His coordination, making this model may be a valuable system in the development of electrochemical biosensors. The ability to use Mimochrome IV and its analogues in biosensor technology has been explored in different works, that are aimed at use these artificial molecules as bio-recognition elements/signal transducers for biosensing.

Therefore, we undertook the NMR analysis of Fe(III)-Mimochrome IV. A preliminary analysis was carried out in aqueous solution (phosphate buffer 10 mM, pH = 6.5). Although broad resonances were observed, four signals (at 21.5 ppm, 19.9 ppm, 14.5 ppm and 10.7 ppm) could be reasonably assigned to the heme methyl protons. The spread of these resonances is 10.8 ppm, still indicative of a preferred orientation of axial ligands (~5 ppm is the spread observed for freely rotating axial ligands) (**Fig. 62**).

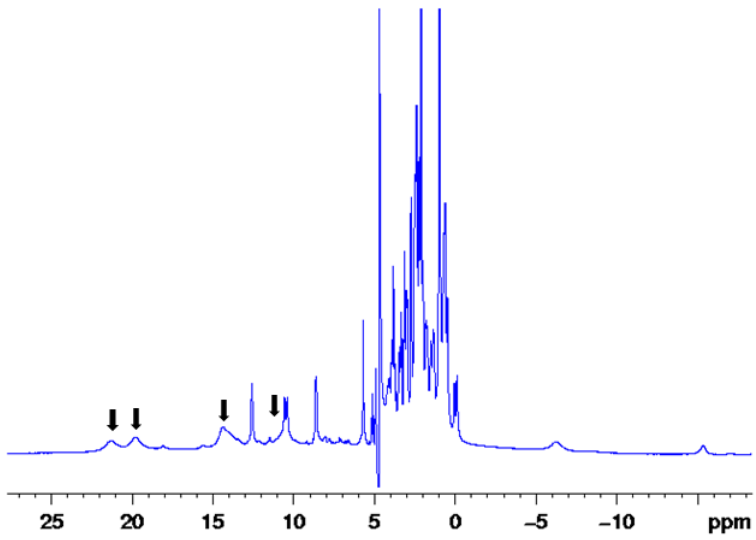


Fig 62. Fe(III)-Mimochrome IV 1D ^{1}H spectrum in phosphate buffer (10 mM, pH 6.5).

However, in order to improve the NMR spectral quality, the analysis was performed in phosphate buffer (10 mM, pH 6.5)/TFE/DMSO solution (60/20/20, v/v/v). In particular, TFE is a well-known helix-inducing solvent, and the addition of a small amount of TFE to the Fe(III)–Mimochrome IV solution resulted in an increase of the helical content. The folding induced by TFE causes the two helical peptides to enwrap the heme faces, thus positioning the imidazole rings appropriately with respect to the porphyrin

ring. The correct folding in the mixed water–TFE solvent is relevant for catalytic Mimochromes. In fact, the presence of an organic co-solvent does not inactivate this class of artificial enzymes, as usually observed for natural enzymes. This finding is very significant for the catalysis of substrates with limited aqueous solubility. Also, the addition of DMSO, a high viscosity solvent (Z = 2 mPa s at T = 298 K), to an aqueous solution, increases the viscosity of the medium and modifies the relaxation rates, thus affecting the spectral lines.

The 1D spectrum is typical of low-spin S = ½ bis-His ferriheme-proteins. It shows several signals shifted out of the diamagnetic region, as a consequence of the effect of the paramagnetic iron (III) ion. In particular, four resonances in the region spanning from 5 to 30 ppm can be attributed to the heme methyl groups (**Fig. 63**). The methyl prtons have a mean shift of 18.5 ppm and a spread of ca. 20 ppm, indicative of fixed orientation of the heme axial ligands. This is typical of heme-proteins, in which the axial ligands are held in fixed orientation by interaction with the protein matrix. On the contrary, small heme protein models usually have axial ligands that rotate very rapidly (on the order of hundreds of thousands to a million times per second), and an average hyperfine shift is detected for each heme substituent.[154,155] In Mimochrome IV, the presence of the peptide chains constrains the histidines to be uniquely oriented.

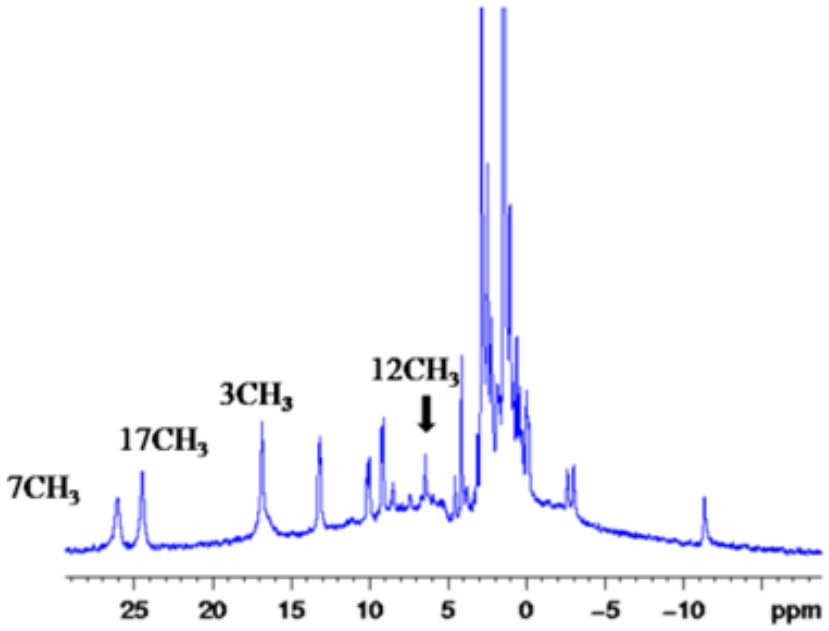


Fig. 63 Fe(III)-Mimochrome IV 1D ^{1}H spectrum.

The assignment of the proton methyl resonances was performed starting from the 5-meso proton at -1.06 ppm, since is the only proton in dipolar contact with the two methyl groups, 7-CH3 and 3-CH3 at 26.1 ppm and 16.9 ppm, respectively. The methyl resonance at 24.5 ppm exhibits dipolar connectivities to a signal at -0.77 (15-meso) and to two protons at about 3 ppm (reasonably attributed to the propionic methylene protons), and therefore it was attributed to the methyl groups at position 17. Finally, the resonance at 6.67 was attributed to methyl at position 12 (**Fig. 64**).

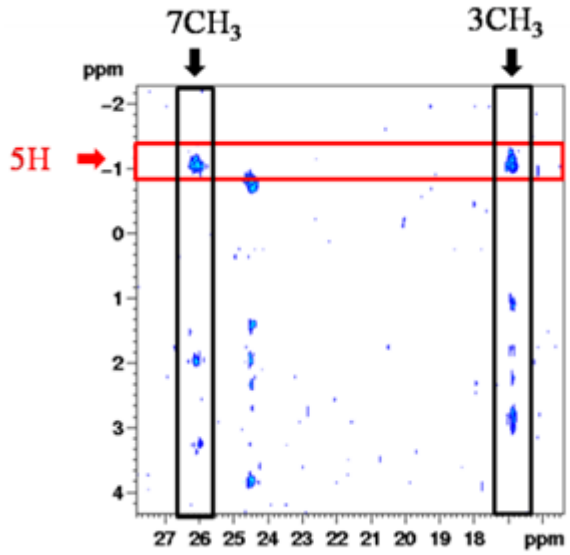


Fig. 64 Fe(III)-Mimochrome IV details of the NOESY spectrum.

In order to determine the axial ligand orientations of Fe(III)–Mimochrome IV, the methyl chemical shifts measured at 298 K were introduced into the empirical equation proposed by Turner:

$$\delta_i(\text{ppm}) = \cos\beta[38.0\ \sin^2(\theta_i-\phi)-4.1\ \cos^2(\theta_i+\phi)-15.9]+13.8 \qquad \text{Eq. (1)}$$

where δ_i is the hyperfine shift of the *i*th methyl, θ_i is the angle between the metal-*i*th methyl direction and the metal-N_{23} nitrogen axis, β is the acute angle between the two histidine planes, and ϕ is the angle between the bisector of the angle β and the metal-N23 direction (see **Fig. 65** for nomenclature). Eqn. (1) provided β and ϕ values of 51° and 34°, respectively.

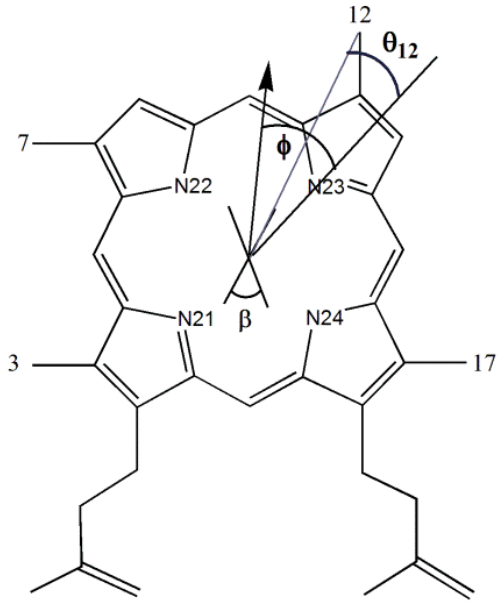


Fig. 65 Schematic representation of the heme moiety. The reference axis is taken along the metal- N23 direction. The β angle defines the acute angle between the two histidine planes and the ϕ angle defines the average orientation of the His planes, projected on the heme plane, with respect to the N_{21}-N_{23} direction.

A comparison between the experimental and calculated values of ^1H methyl chemical shifts is schematically illustrated in **Fig. 66**, which shows the good agreement between the pattern of distribution of the experimental shifts and the theoretical fit. Table 9 reports the β and ϕ values calculated for the Fe(III)–Mimochrome IV complex compared with those observed in the Co(III)–Mimochrome IV crystal and NMR structures.

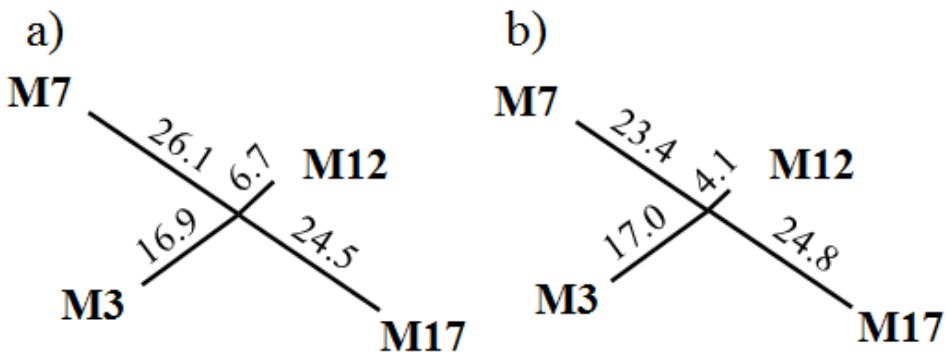


Fig. 66 Diagram showing **a)** experimental and **b)** calculated chemical shifts of the heme methyls. Lines have the iron at the origin and are directed towards each substituent, with a length proportional to the shift. The implicit orientation of the heme is as represented in **Fig. 65**.

Table 9 Comparison of β and ϕ in Mimochrome IV complexes

	β (°)	ϕ (°)
Fe(III)- Mimochrome IV	51	3
Co(III)- Mimochrome IV X-ray	57	47
Co(III)- Mimochrome IV NMR	52	33

The orientation of axial histidines in Fe(III)–Mimochrome IV is nearly identical to that observed in the solution structure of its cobalt complex, with the imidazole planes nearly eclipsing the Np–Co– Np bonds. In contrast, the metal binding site shows a different geometry in solid state, where the histidine planes are almost perpendicular. These data highlight that the main issue in determining the histidine orientation in Mimochrome IV is the direction of the two helices with respect to the porphyrin plane, which, in the Co(III)– Mimochrome IV crystal is affected by crystal packing interactions.

Conclusions and perspectives

5. Conclusions and Future perspectives.

The design of metalloprotein models, tailored to specific applications, is emerging as an important topic in protein chemistry. To date, several groups have made significant progress toward the design of proteins with functions and activities similar to those of natural enzymes. Metalloprotein mimetics have been developed through the introduction of novel metal-binding sites within the naturally occurring proteins as well as through de novo design. Such studies are stimulated by the ambition to shed light on the fundamental features for catalytic activity, and by the possibility to develop novel synthetic proteins for biotechnological applications. Tailoring synthetic models requires the development of sophisticated molecular architectures that distil the quintessential elements responsible for activities. Thus peptide-based metalloprotein models seem to be suitable candidates to mimic both the structural features and reactivity of the natural systems.

This approach has been successfully applied to different natural proteins involved in biological electron transfer, catalysis and gene regulation. A protein family that has received considerable attention is the heme-protein family. Heme-proteins are involved in a wide range of biological processes, and over the years, a large number of peptide and protein-based heme-protein mimetics have been developed, in order to determine how the protein matrix tunes the properties of the heme to evoke the wide variety of activities.

The aim of this PhD thesis was to engineer new heme-protein mimetics characterized by a well-defined three-dimensional structure, thus allowing structural characterization. In particular, the research activity was focused on the rational design, synthesis, and characterization of new Mimochrome

analogues. These molecules belong to a class of heme-proteins developed by the *Artificial Metallo Enzyme Group* of Naples.

In this PhD project, two different strategies have been followed to take into account all the variables necessary to specify both the structure and the function in the Mimochrome family of molecules. In particular, new mimetics have been designed by:

- incorporation of non-coded amino acid with α-helix stabilizing effect into the Mimochrome VI structure;
- elongation of the peptide chains, that allows the introduction of a large number of intra-chain and inter-chain interactions.

Both strategies afforded to the design and synthesis of two new molecules, named Mimochrome VI-2U1L and Mimochrome VII.

During the research activity, the important goal of the structural characterization of the Co(III)- complexes of these analogues was reached. The information derived on Mimochrome VI-2U1L structures will be used to obtain a detailed structure-activity relationship analysis on Mimochrome VI and its analogues. In addition, the availability of the solution structure of Mimochrome VII should allow to engineer new molecules. Indeed, in this thesis the design of an asymmetrical penta-coordinated model, Mimochrome VII asym derivative, characterized by a distal cavity that contains an arginine residue, mimicking Arg^{38} of Horseradish Peroxidase (HRP), and a non-coordinating His residue to mimic His^{42} of HRP was done, and its characterization is now under course.

With the aim to fully characterize the active site structures of Mimochromes, another important results of this thesis work is the NMR characterization of the paramagnetic peptide-based hemeprotein mimetics, Fe(III)–Mimochrome IV complex. By combining an empirical equation, proposed for bis histidine ferriheme-proteins, with experimental data obtained for Fe(III)–Mimochrome IV, information about histidine orientations were obtained. In summary, it has been shown that it is possible to determine the histidine orientations in artificial peptide-based heme-protein models by taking advantage of a methodology previously used for natural ferriheme-proteins.

The results herein reported highlight that in small models, which lack the complex architecture of the natural protein, the carefully engineering of numerous and diverse interactions, such as hydrophobic, ion-pair and coordination interactions, is crucial for the development of molecules with unique structures and functions. In addition, an important result of this thesis has been the successfully application of an empirical equation, describing the relationship between the porphyrin methyl hyperfine shifts and the position of the axial ligand(s), to these artificial heme-protein.

In conclusion, the different design strategies and experimental methodologies reported in this work will in future allows to perform iterative process of redesign and characterization, with the final goal of obtaining functionally active molecules with a well defined structure.

Experimental part

6. Mimochrome VI-2U1L

6.1. General

All 9-fluorenylmethoxycarbonyl (Fmoc) protected amino acids, Nova Syn TG Sieber resin and coupling reagents: N-hydroxybenzotriazole (HOBt), *O*-7-Azabenzotriazol-1-yl-N,N,N',N''-tetramethyluronium hexafluorophosphate (HATU) and benzotriazole-1-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) were purchased from NovaBiochem (EMD Biosciences, La Jolla, CA). Silica gel 60 (230-400 mesh) was from Merck. Precoated silica G-60 plates, F254, were used for thin-layer chromatography (tlc). All solvents used in the peptide synthesis and purification were anhydrous and HPLC grade respectively, and were supplied by Romil. Piperidine and scavengers (ethanedithiol, triisopropylsilane) were from Fluka. N,N-Diisopropylethylamine (DIEA), trifluoroacetic acid (TFA) were supplied from Applied Biosystems. N,N-dimethylformamide (DMF), dichloromethane (DCM), pyridine, ethanol, methanol and 1-methyl-2-pyrrolidone (NMP) were supplied by Romil. Deuteroporphyrin IX was from Porphyrin Products. Cobalt(II) acetate was purchased from Sigma Aldrich. Solvent mixtures are indicated in the respective sections.

The structures were generated with Visual Molecular Dynamics (VMD; http://www.ks.uiuc.edu/Research/vmd/)[156] and Accelrys Discovery Studio 3.0 (http://accelrys.com/).[157]

6.2. Synthetic strategy

Mimochrome VI-2U1L was prepared similarly to other Mimochromes.[118] The two peptide chains were synthesized by the solid-phase method using the 9-fluorenylmethoxycarbonyl (Fmoc) protection

strategy. After selective deprotection of Lys[9] residue on the two chains, the peptides were cleaved from the resin, without removing the other side chain protecting groups (**Fig.67**). They were then coupled to deuteroporphyrin IX in solution to afford, after side chain deprotection, Mimochrome VI-2U1L.

6.2.1. Peptides synthesis

Mimochrome VI-2U1L decapeptide and Mimochrome VI-2U1L tetradecapeptide were synthesized using Fmoc (9-fluorenylmethoxycarbonyl) solid phase peptide synthesis (SPPS) on an ABI 433 automatic peptide synthesizer on a 0.2 mmol scale. The resin used for the peptides was the super acid labile Nova Syn TG Sieber (substitution level 0.2 mmol/g) (**Fig. 64**).

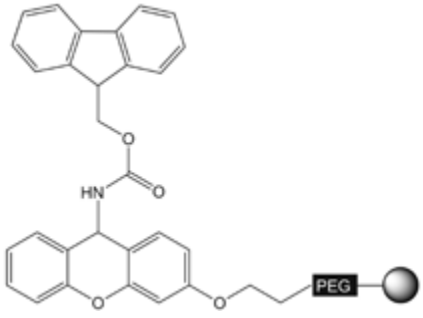


Fig. 67 Nova Syn TG Sieber resin

For all the peptides the same synthetic procedures were used. The first step was the synthesis of the fully protected peptides. The synthetic procedure can be summarized as follow. The α- Fmoc group, was removed at the beginning of every cycle with a 20% piperidine solution. After deprotection, the resin was washed with NMP to remove the piperidine. The peptide resin was then ready for coupling. All amino acids were coupled using a 0.45 M HBTU/HOBT solution in DMF. In the coupling step, the activated Fmoc amino acid reacts with the amino-terminal group of the growing peptide chain to form a peptide bond.

Deprotection and coupling steps were repeated with each subsequent amino acid, until the chain assembly was completed. When the coupling was

complete, the resin was washed with NMP. Peptides N-terminal amino groups were acetylated with acid acetic (1mmol, 57 µl).
For the synthesis of the two peptides the following amino acids were used:

Mimochrome VI-2U1L decapeptide: Fmoc-Asp(OtBu)-OH; Fmoc-Glu(OtBu)-OH; Fmoc-Gln(Trt)-OH;Fmoc-Aib-OH; Fmoc-Leu-OH; Fmoc-Ser(tBu)-OH; Fmoc-Lys(Mmt)-OH; Fmoc-Arg (Pbf)-OH.

Mimochrome VI-2U1L tetradecapeptide: Fmoc-Asp(OtBu)-OH; Fmoc-Gln(Trt)-OH; Fmoc-Leu-OH; Fmoc-His(Trt)-OH; Fmoc-Ser(tBu)-OH; Fmoc-Lys(Mmt)-OH; Fmoc-Arg(Pbf)-OH; Fmoc-Lys(Boc)-OH; Fmoc-Ile-OH; Thr(tBu)-OH.

6.2.2. Mmt deprotection and cleavage of Mimochromes peptides

Fully protected peptides amide, except for the Lys^9, were generated from the resin separately, following the same protocol. (**Fig. 66**) The protecting group on N-ε function of the Lys^9 was chosen such as since its cleavage conditions were different than those used for the other side chain protecting groups. Specifically, The Mmt protecting group was used, since it can be easily removed by repeated treatments with a solution containing 10% acetic acid and 20% trifluoroethanol (TFE) (v/v) in CH_2Cl_2, without affecting the other protecting groups. This treatments was performed at the end of the synthesis, when the peptide is still linked to the resin and was repeated until no yellow or red trityl cations were detected in the eluent.

After this step, cleavage of the peptides from the resin was achieved with a solution 1%TFA in DCM. 3-4 resin volumes of a freshly prepared cleavage mixture (1%TFA in DCM, 20 ml) were added. The acidic mixture

was incubated for 2 minutes, under mixing. The solution was filtered with a vacuum pump, into an ice-cold flask containing 2 ml of 10% pyridine/methanol (v/v).

The reaction was followed by TLC (Thin Layer Chromatography, chloroform/methanol/acetic acid 80:18:2); each step was repeated until no product was detected in the collected fractions. The fractions containing the product were combined and evaporated under reduced pressure up to 5% of the volume. Ice-cold water was added to the residue and the mixture was cooled on ice to aid precipitation of the protected peptides. The product was filtered, washed several times with fresh water, and dried under vacuum to give the crude C-terminal peptides amide. Both peptides were homogeneous as determined by analytical HPLC and by ESI –MS.

Fig. 68 Protocol of Mmt deprotection and cleavage from the resin.

6.2.3. Synthesis of the monosubstituted intermediate decapeptide-deuteroporphyrin IX.

A suitable procedure for the coupling reaction between the decapeptide and the deuteroporphyrin IX was adopted, to avoid the formation of the disubstituted intermediate. In particular, the decapeptide solution (419 mg, 0.20 mM , M.W. = 2119.6) was slowly added to the

solution containing deuteroporphyrin IX (1.5 eq., 0.30 mol, 173 mg) and HATU (activating reagent, 1eq., 75.1 mg). the pH was adjusted with DIEA, 7 eq. (pH ≈ 8.0). In this way, the porphyrin was never in the presence of peptide excess, thus favoring monosubstitution. The reaction mixture was stirred for 2 h at room temperature, and the outcome was followed by analytical TLC, solvent system chloroform/methanol 90/10.

The crude product was purified by Flash Chromatography using a Biotage Isoelera flash purification system, equipped with a diode-array detector, on SNAP HP 100 g silica column (5 x 30cm), using a chloroform/methanol gradient from 0 to 15% methanol. The desired protected monoadduct was separated from the porphyrin and the un-reacted peptide, and analyzed by RP-HPLC, on a C-8 column, using a linear gradient elution of water/acetonitrile 0.1% TFA. The chromatographic profile, shown in figure , evidences two peaks of similar intensity at 45.7 and 46.7 minutes **(Fig. 69**).

The ESI-MS analysis of the two peaks reveals the same mass spectra, and in both mass spectra in addition to the main peak of 1307.35, corresponding to the ion $[M-2H]^{2+/}2$ is observed. Attempt to separate the two peaks through Flash Chromatography were not successful, due to their similar Rf.

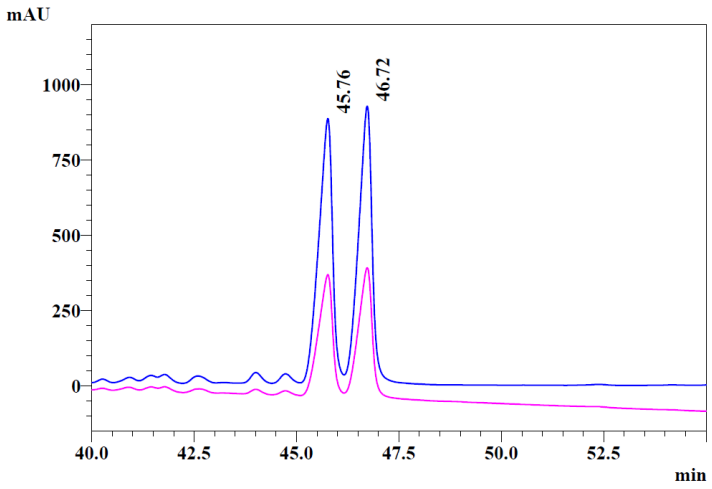


Fig. 69 HPLC chromatogram and ESI-MS spectrum of monosubstituted intermediate decapeptide-deuteroporphyrin IX.

6.2.4. Synthesis of disubstituted intermediate decapeptide-DPIX-tetradecapeptide

The Mimochrome VI-2U1L apo form was synthesized by coupling the tetradecapeptide (573 mg, 0.18 mmol, P.M. = 3239 mg/mmol with the monosubstituted decapeptide-deuteroporphyrin IX intermediate (357 mg, 0.14 mol, P.M. = 2612mg/mmol). A small excess of tetradecapeptide was used, to drive the reaction to completion.

The outcome of the reaction was followed by analytical TLC. Subsequently, the desired product was separated from the monosubstituted intermediate and the un-reacted tetradecapeptide by Flash Chromatography using a linear gradient elution of $CHCl_3/CH_3OH$.

The protected apo form was than deprotected using a cleavage mixture of 94% TFA, 2.5% EDT, 1% TIS, 2.5% H2O v/v/v/v; 10 ml. The reaction was conducted at 273 K for 1h and at room temperature for the second hour, with stirring; the final mixture was concentrated on a rotary evaporator to a

volume of approximately 1-2 mL. Extraction of the scavengers and precipitation of the crude product was achieved by addition of cold diethyl ether. The mixture was centrifuged, the supernatant was removed and the precipitate was washed twice with three volumes of cold diethyl ether. The mixture was dried to remove diethyl ether, redissolved in water 0.1 % TFA and analyzed by LC-MS (Vydac C-18 column, using a gradient of acetonitrile in 0.1% aqueous TFA, 5 % to 80% over 35 min). The crude material was then purified by preparative RP-HPLC (Vydac 2.2 cm C18 column at 22 mL min-1, using a gradient of acetonitrile in 0.1% aqueous TFA, 10% to 80% over 58.4 min); the pooled fractions containing the desired product were lyophilized. The chromatogram showed the presence of two peaks that have same area, slightly different retention times and the same mass spectrum.

6.2.5. Cobalt insertion

Cobalt ion was inserted into Mimochrome VI-2U1L apo form (13 mg, $3.8 * 10^{-3}$ mmol) according to literature procedure.[45] Co(II) acetate (10 molar excess) was added to a solution of pure Mimochrome VI-2U1L apo form in 2/3 TFE/AcOH (v/v) (19 mL, final concentration, $2.0*10^{-4}$ M), and the reaction mixture was kept at 348K for 2 h, refluxing under nitrogen. The reaction was monitored by analytical HPLC. The solvent was then removed under vacuum, and the product was purified to by preparative RP-HPLC.

6.3. NMR spectroscopy

6.3.1. NMR data collection

Samples for NMR spectroscopy were prepared using by dissolving weighted amounts of Co(III)-Mimochrome VI-2U1L in H_2O/ TFE 60/40

(v/v) V = 0.600 ml). The final concentrations of the samples were between 2 $*10^{-4}$ M and 5 $*10^{-4}$.

All NMR spectra were acquired at at 298 K on a Bruker Avance 600 spectrometer equipped with a triple resonance cryo-probe. Suppression of the water signal was accomplished by excitation sculpting sequence. Two dimensional NOESY, TOCSY and DQF-COSY spectra were acquired at very high digital resolution, using standard pulse sequences. For peak a NOESY spectra were acquired using mixing time ranging from 200 ms and 300 ms. Subsequently, 200 ms mixing time was selected to acquire the NOESY spectra of all peak. TOCSY experiment was performed with spin lock applied for 70 ms.

6.3.2. Resonances assignment and strcutural calculation

NMR Spectra were processed using Bruker TOPSPIN software and analyzed with the CARA program. The ^{1}H chemical shifts (in ppm) of Co(III)-Mimochrome VI-2U1L peak a and Co(III)-Mimochrome VI-2U1L are reported in table 10 and 11 for **peak a** and **peak b** respectively.

Table 10 Co(III)-Mimochrome VI-2U1L **peak a** assignment						
Tetradecapeptide						
Aa	**N-H**	**α -H**	**H-β**	**H-γ**	**H-δ**	**others**
ACE		1.99				
Asp[1]	7.82	4.59	2.81			
Leu[2]	8.16	3.33	1.41 0.53	1.18	0.91 0.70	
Gln[3]	8.30	3.28	1.88 1.79	2.07		$H_{\varepsilon 1,2}$ 7.13 6.49
Gln[4]	7.86	3.68	1.91 1.90	2.26 2.21		$H_{\varepsilon 1,2}$ 7.20 6.53
Leu[5]	7.61	3.55	0.95 0.50	0.81	0.62 0.17	
His[6]	6.31	2.33	1.50 1.11		$H_{\delta}1$ 9.23 $H_{\delta}2$ -0.49	$H_{\varepsilon 1}$ -0.16
Ser[7]	7.47	3.83	3.75			
Gln[8]	7.60	4.13	2.12 1.93	2.29		$H_{\varepsilon 1,2}$ 7.02 6.59
Lys[9]	7.35	3.98	1.68	1.53	1.44	$H_{\varepsilon,\varepsilon'}$ 3.13 2.43 $H\zeta$ 8.23
Arg[10]	8.02	4.11	1.85 1.73	1.55 1.44	3.14 3.05	H_{ε} 7.05
Lys[11]	7.82	4.13	1.95 1.93	1.55 1.42	1.70	$H_{\varepsilon,\varepsilon'}$ 2.96
Ile[12]	7.95	4.06	2.05	$H\gamma_{12,13}$- CH_2 1.71 1.31 $H\gamma 2$- CH_3 1.05	0.96	
Thr[13]	7.95	4.37	4.40	1.45		
Leu[14]	7.95	4.42	1.85	1.68	0.97	
NH_2	7.25, 6.93					

Decapeptide						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
ACE		2.01				
Asp[1]	8.02	4.62	2.86 2.81			
Glu[2]	8.33	3.99	2.01 1.97	2.39		
Aib[3]	8.02		1.41 1.29			
Gln[4]	7.95	3.85	1.97 1.90	2.33 2.32		$H_{\varepsilon 1,2}$ 7.24 6.56
Leu[5]	7.80	3.96	1.64 1.57	1.39	0.84 0.77	
Ser[6]	7.59	3.85	3.60			
Aib[7]	7.58		1.38 1.28			
Glu[8]	7.68	3.88	1.92 1.85	2.24		
Lys[9]	7.55	3.65	1.43 1.29	0.97 0.87	0.96	$H_{\varepsilon,\varepsilon'}$ 2.82 H_{ζ} 7.41
Arg[10]	7.56	4.14	1.91 1.74	1.64 1.59	3.15	Hε 7.13
NH_2	7.08 6.85					

Deuteroporphyrin IX			
2 -α,α'CH_2	4.93, 4.39	8H	9.48
18-α,α'CH_2	4.96, 4.27	10 H	10.60
2 -β,β'CH_2	3.45, 3.02	12CH3	3.86
18-β,β'CH_2	3.14, 3.07	13 H	9.40
3CH3	3.79	15 H	10.28
5H	10.50	17CH3	3.68
7 CH3	3.79	20 H	10.59

Table 11: Co(III)-Mimochrome VI-2U1L **peak b** assignment						
Tetradecapeptide						
Aa	**H-N**	**H- α**	**H-β**	**H-γ**	**H-δ**	**others**
ACE		1.99				
Asp[1]	7.78	4.58	2.79			
Leu[2]	8.16	3.35	1.43 0.50	1.24	0.97 0.70	
Gln[3]	8.30	3.28	1.89 1.80	2.06		
Gln[4]	7.89	3.69	1.90	2.26 2.21		
Leu[5]	7.63	3.55	1.04 0.54	0.82	0.60 0.05	
His[6]	6.35	2.32	1.50 1.13		$H_{\delta1}$ 9.24 $H_{\delta2}$ -0.43	$H_{\varepsilon1}$ -0.174
Ser[7]	7.50	3.81	3.74			
Gln[8]	7.68	4.07	2.07 1.95	2.29		
Lys[9]	7.43	3.96	1.63	1.47	1.62	$H_{\varepsilon,\varepsilon'}$ 3.15 2.30 H_{ζ} 8.39
Arg[10]	8.03	4.05	1.83 1.72	1.50	3.12 2.99	Hε 6.99
Lys[11]	7.80	4.09	1.95	1.55 1.41	1.67	$H_{\varepsilon,\varepsilon'}$ 2.94
Ile[12]	7.97	4.04	2.05	$H\gamma_{12,13}$- CH_2 1.71 1.30 Hγ2- CH_3 1.04	0.94	
Thr[13]	7.95	4.34	4.41	1.46		
Leu[14]	7.97	4.42	1.86	1.68	0.97	
NH_2						

Decapeptide						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
ACE		2.02				
Asp[1]	8.00	4.63	2.86 2.81			
Glu[2]	8.35	3.99	2.03 1.99	2.39		
Aib[3]	8.03		1.42 1.32			
Gln[4]	7.95	3.83	1.97	2.31		-
Leu[5]	7.81	3.96	1.41	1.65	0.85 079	
Ser[6]	7.61	3.84	3.64			
Aib[7]	7.55		1.38 1.25			
Glu[8]	7.67	3.87	1.85	2.23		
Lys[9]	7.52	3.53	1.40	1.23	0.85	Hε,ε' 2.80
Arg[10]	7.50	4.12	1.91 1.73	1.62	3.15	Hε 7.13
NH_2						

Deuteroporphyrin IX			
2 -α,α'CH_2	4.95, 4.28	8H	9.54
18-α,α'CH_2	4.95, 4.41	10 H	10.63
2 -β,β'CH_2	3.14, 3.05	12CH3	3.86
18-β,β'CH_2	3.46, 3.01	13 H	9.40
3CH3	3.71	15 H	10.42
5H	10.45	17CH3	3.72
7 CH3	3.82	20 H	10.60

Intensities of dipolar connectivities in the two-dimensional NOESY spectrum, obtained using a 200 ms mixing time, were measured using the integration subroutine of the CARA program. Peak volumes were converted into upper limit of inter-atomic distances, to be used as input for structural calculation.

A total of 395 and 341 NOESY cross peaks was assigned, integrated and transformed in upper distance limits with the program CALIBA,[158] for peak a and peak b, respectively. Of these NOE upper distance limits, 142 for peak a and 121 for peak b were removed, because they do not effectively restrict the conformation. In conclusion, we obtain two set of 253 (196 intra-chain distances and 57 distances involving and 255 upper distance limit restraints that were used as input for structural calculation.

In order to perform the simulated annealing, two non standard residue were added to the CYANA standard library: a residue containing Co(III)-deuteroporphyrin IX, that is anchored to the peptide chain through the use of special linker residues (as described in CYANA tutorial) and a modified lysine residue, which has only one proton on the side chain N^{ζ} atom, since this atom is involved in formation of amidic bond with one of the deuteroporphyrin propionyl group.

Additional restraints were added to simulate the histidine coordination, and the amide bond formation between the lysine side chain and the deuterporphyrin. The quality of the obtained structures was evaluated in terms of deviations from ideal lengths and bond angles through a Ramachandran plot obtained using the program PROCHECK-NMR.

7. Mimochrome VII

7.1. Synthetic strategy

Mimochrome VII was prepared similarly to Co(III)-Mimochrome VI-2U1L; the peptides were synthesized by the solid-phase method using the 9-fluorenylmethoxycarbonyl (Fmoc) protection strategy. The resin used for the was Nova Syn TG Sieber (9-fmoc-amino-xanthen-3-yloxy TG resin, substitution level 0.2 mmol/g). For the synthesis of the peptide the following amino acids were used:

Ac-Thr^{1}(tBu)-Asn^{2}(Trt)-Asp^{3}(tBuO)-Trp^{4}(Boc)-Asp^{5}(tBuO)-Pro^{6}-Thr^{7}(tBu)-His^{8}(Trt)-Gln^{9}(Trt)-Glu^{10}(tBuO)-Lys^{11}(Mmt)-Leu^{12}-Glu^{13}(tBuO)-Lys^{14}(Boc)-Arg^{15}(Pbf)-Leu^{16}-Lys^{17}(Boc)-NH_2

After selective deprotection of Lys^{11} residue, the peptid was cleaved from the resin, without removing the other side chain protecting groups. There were then coupled to deuteroporphyrin IX in solution to afford, after side chain deprotection, Mimochrome VII apo form.

7.1.1. Synthesis of the deuteroporphyrin mono-substituted derivative

The first step was the deuteroporphyrin activation:

- the deuteroporphyrin IX dihydrochloride (44 mg, 0.075 mmol, M.W. 583.5 g/mol) was dissolved in 1 mL DMF and 43.5 µL DIEA (5 eq., 0.250 mmol, M.W., 129.24 g/mol; density, 0.742 g/mL).

- 1 equivalents of HATU (19 mg, M.W., 380.23 g/mol) was dissolved in 500 µL DMF, and added to the deuteroporphyrin.

Subsequently, the eptapeptide (190 mg, 0.0500 mmol) was dissolved in 1 mL DMF and DIEA 17.e µL (2 eq.). The peptide solution was then added dropwise to the deuteroporphyrin solution.

The pH was adjusted with DIEA (pH ≈8.0). the reaction mixture was stirred at room temperature, and monitored by tlc (solvent system chloroform/methanol 90:10). After one hour and a half, the reaction mixture was evaporated under reduced pressure and dried. The crude product was purified on a silica gel column (30 g, 230–400 Mesh, 3.5 x 60 cm), with stepwise elution using a chloroform/methanol gradient from 0 to 10% methanol. The product was eluted at 10% methanol, evaporated under reduced pressure and dried.

7.1.2. Synthesis of the apo form

The monopeptide adduct (80 mg, 0.019 mmol) was dissolved in 1 mL of DMF. HATU (8.6 mg, 0.023 mmol) was dissolved in 200 µL of DMF and added to the monopeptide solution. 6 eq. (19.8 µL) of DIEA were added.

The eptapeptide was dissolved in 900 µL of DMF. This solution was then added to the monopeptide adduct solution and the reaction was allowed to proceed at room temperature for hour and a half and the pH was checked during the reaction time. The reaction was monitored by tlc (solvent system chloroform/methanol 90:10). After one hour and a half the reaction mixture was evaporated under reduced pressure. The product was dissolved in chloroform, and extracted six times with water, and evaporated again.

Side chain deprotection was achieved by addition of the cleavage mixture (thioanisole/H_2O/EDT/TFA 0.25/0.5/0.5/8.75, v/v/v/v) (1,2-

ethanedithiol: EDT) at 273 K for 2.5 h. The reaction mixture was concentrated on a rotary evaporator to a volume of approximately 1-2 mL. Extraction of the scavengers and precipitation of the crude product was achieved by addition of cold diethylether. After six washing with diethylether, the crude material was then dried in vacuo and analyzed by RP-HPLC (Vydac C18 column, using a H_2O 0.1% trifluoroacetic acid (TFA) (v/v) (solvent A) and acetonitrile 0.1% TFA (solvent B) linear gradient, from 5% to 70% (solvent B) over 60 min, flow rate 0.5 mL min^{-1}). ESI-MS confirmed the expected molecular weight 4833.3 amu) for the most abundant product (**Fig 70**).

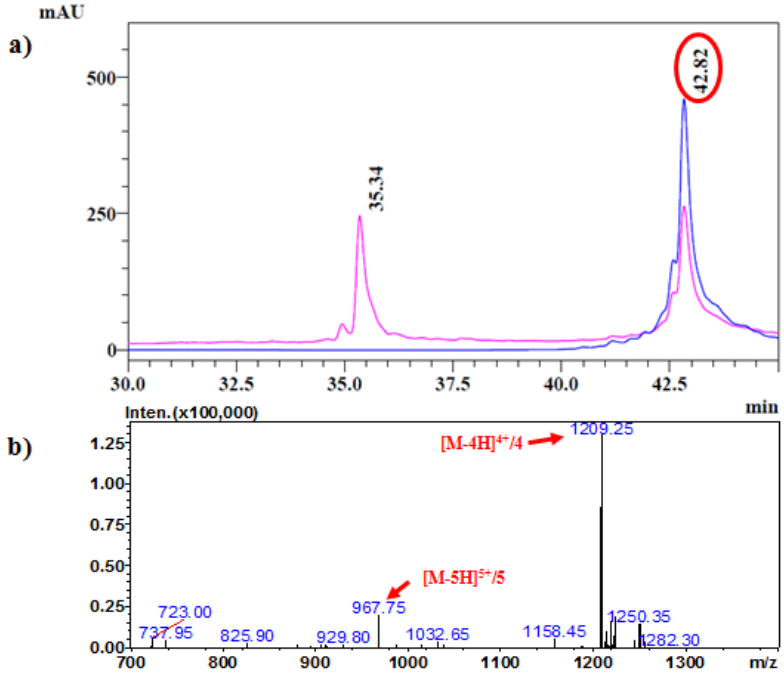


Fig. 70 HPLC chromatogram and ESI-MS spectrum of symmetric Mimochrome VII apo form.

7.1.3. Iron insertion

Iron was inserted into Mimochrome VII symmetric according to the acetate method.[45]

Iron (II) acetate (10 molar excess), was added to a solution of pure free base in TFE/Acetic acid (6/4 v/v), and the reaction mixture was kept at ≈ 358 K for 1 h, refluxing under nitrogen. After 20 minutes, the reaction mixture started to turn slowly brown. After 1h, the HPLC chromatogram showed the presence of a new peak with the characteristic UV-visible spectrum of the iron (III) deuterophorphyrins at acidic pH, λmax = 388 nm The product was then purified by preparative RP-HPLC, and the LC-MS IT-TOF (ion trap (IT) and time-of-flight (TOF) mass spectrometry) analysis of the pure product confirmed the identity of the desired compounds (expected molecular weight 4887.2 amu) (**Fig. 71**).

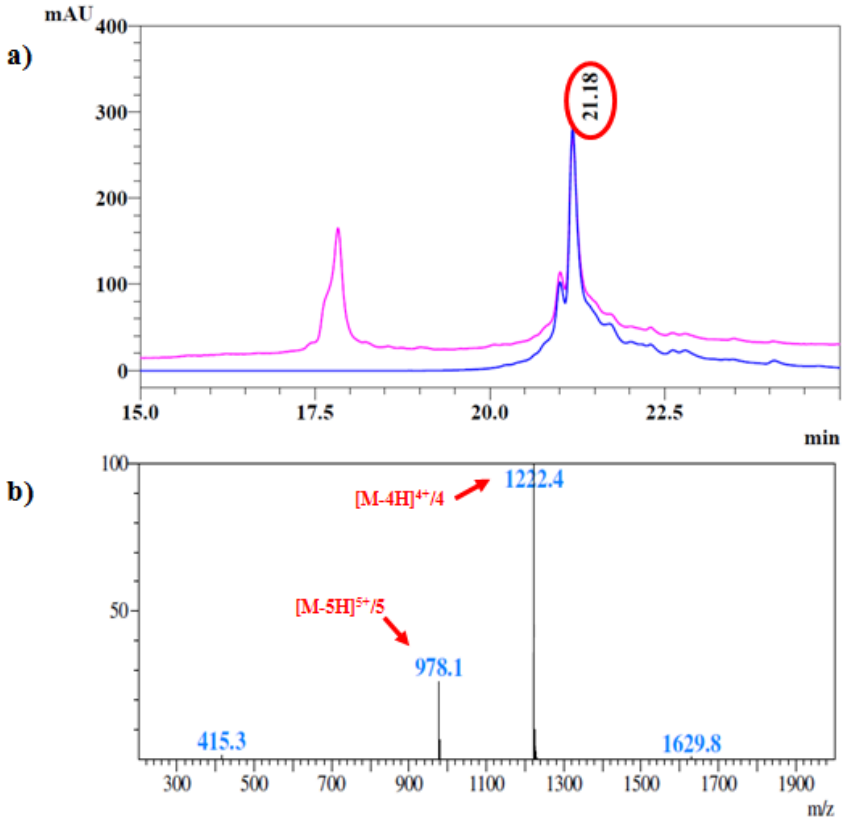


Fig. 71 HPLC chromatogram (a) and LC-MS IT-TOF (b) spectrum of Fe(III)-Mimochrome VII.

7.1.4. Cobalt insertion

Cobalt was inserted into Mimochrome VII according to the acetate method.[45] Co(II) acetate (10 molar excess) was added to a solution of pure free base in TFE/Acetic acid (6/4 v/v), and the reaction mixture was kept at 348 K for 1 h, refluxing under nitrogen. After 20 minutes, the reaction mixture started to turn slowly brown. After 1h, the HPLC chromatogram showed the presence of a new peak with the characteristic UV-visible spectrum of the cobalt(III) deuterophorphyrins at acidic pH, λmax = 416 nm.

The product was then purified by preparative RP-HPLC (**Fig.72**), and the ESI-MS IT- analysis of the pure prodcut confirmed the identity of the desired compound (expected molecular weight 4890.3 amu).

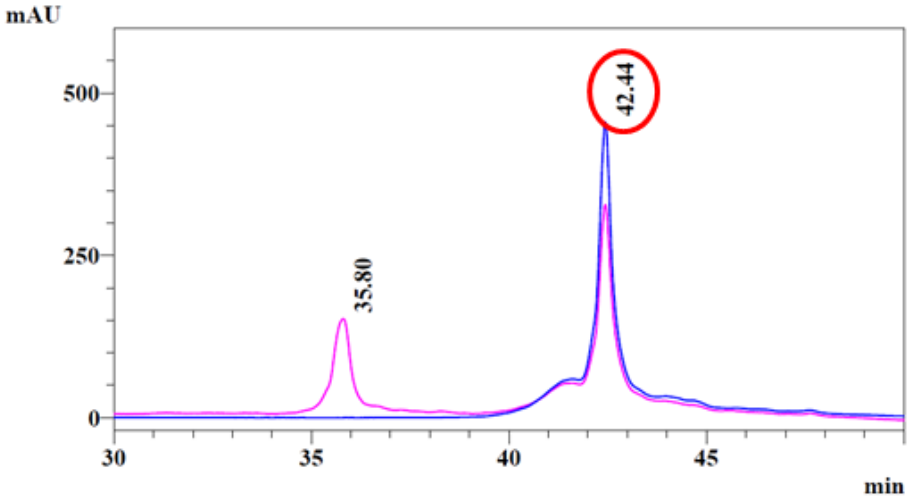


Fig. 72 HPLC chromatogram of Co(III)- Mimochrome VII

7.2. CD and UV-Vis spectroscopy

CD spectra were obtained at 298 K on a Jasco J-715 dichrograph. Data were collected from 260 to 190 nm at 0.2 nm intervals with a 20 nm min^{-1} scan speed, at 1 nm bandwidth and at 16 s response. Cuvette path length of 0.1 cm was used. Sample concentration was $1*10^{-5}$ M. CD intensities are reported as molar ellipticity *per* residue (mdeg cm^2 $dmol^{-1}$ res^{-1}). All data are blank subtracted.

UV-Vis spectra were recorded on a Cary Varian 50 Probe UV Spectrophotometer. All measurements were performed at room temperature. Quartz cuvettes with a path length of 1.0 cm were used for most measurements. Wavelength scans were performed from 200 nm to 700 nm, with a 300 nm min^{-1} scan speed. All data are blank subtracted.

pH titration was performed in aqueous solution, using TFA solutions (0.1 M and 1M) trifluoroacetic acid to lower pH. The concentration of Mimochrome VII was 1.0 * 10-5M. Dilution was less than 3 % and considered in the final data. The pKa's were determined by monitoring the changes in absorbance at 388 nm as a function of proton concentration.

The data were analyzed assuming the following equilibria:

$$[AH_2^{2+}] \overset{Ka_1}{\rightleftharpoons} [AH^+] + [H^+] \overset{Ka_2}{\rightleftharpoons} [A_0] + [H^+]$$

The equilibrium constants can be expressed as:

$$K_{a1} = \frac{[AH^+]\cdot[H^+]}{[AH_2{}^{2+}]} \;;\; K_{a2} = \frac{[A_0]\cdot[H^+]}{[AH^+]} \qquad \textbf{(eq.2)}$$

Each of them can be expressed in dependence of $[A_0]$. The total absorbance and the total concentration can be expressed as sum of the all species:

$$Abs_{tot} = \varepsilon_0 \cdot [A_0] + \varepsilon_1 \cdot \frac{[A_0] \cdot [H^+]}{K_{a2}} + \varepsilon_2 \cdot \frac{[A_0] \cdot [H^+]^2}{K_{a1} \cdot K_{a2}} \quad \textbf{(eq.3)}$$

$$[\ ]_{tot} = [A_0] + [AH^+] + [AH_2^{2+}] \quad \textbf{(eq.4)}$$

Next, $[A_0]$ is isolated from eq. 4, and substituted in eq. 3, obtaining:

$$Abs_{tot} = \frac{[\ ]_{tot} \cdot \left(\varepsilon_0 + \varepsilon_1 \cdot \frac{[H^+]}{K_{a2}} + \varepsilon_2 \cdot \frac{[H^+]^2}{K_{a1} \cdot K_{a2}} \cdot \right)}{1 + \frac{[H^+]}{K_{a2}} + \frac{[H^+]^2}{K_{a1} \cdot K_{a2}}} \quad \textbf{(eq.5)}$$

Multiplying the numerator and the denominator of the right-hand side, by $[\ ]_{tot} \cdot K_{a1} \cdot K_{a2}$, eq. 2 can be written as:

$$\varepsilon_{tot} = \frac{\varepsilon_0 \cdot K_{a1} \cdot K_{a2} + \varepsilon_1 \cdot K_{a1} \cdot [H^+] + \varepsilon_2 \cdot [H^+]^2}{K_{a1} \cdot K_{a2} + K_{a1} \cdot [H^+] + [H^+]^2} \quad \textbf{(eq.6)}$$

where ε_0, ε_1, ε_2 are the molar extinction coefficients of the different species. **Eq.6** was used to fit the observed ε_{388nm} as function of $[H^+]$.

7.3. NMR spectroscopy

7.3.1. NMR data collection

Samples for NMR spectroscopy were prepared using by dissolving weighted amounts of Co(III)-Mimochrome VII symmetric in H_2O/ TFE 50/50 (v/v) (pH = 4.6) (V = 0.500 ml) for a final concentration of $5.3 * 10^{-4}$ mM.

All NMR spectra were acquired at at 298 K on a Bruker Avance 600 spectrometer equipped with a triple resonance cryo-probe. Suppression of the water signal was accomplished by excitation sculpting sequence. Two dimensional NOESY, TOCSY and DQF-COSY spectra were acquired at very high digital resolution, using standard pulse sequences. NOESY spectra were acquired using mixing time between 200 and 350 ms. TOCSY experiment were performed with spin lock applied for 60 or 70 ms.

7.3.2. NMR data analysis

Spectra were processed using Bruker TOPSPIN software and analyzed with the CARA program. The ^1H chemical shifts (in ppm) for c.1 and c.2 are reported in table 12 and 13, respectively.

Table 12: Co(III)-Mimochrome VII **compound 1** assignment.						
Peptide chain linked to propionate at position 2						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
Thr[1]	7.91	4.37	4.27	1.20		
Asn[2]	8.35	4.69	2.65		6.62 7.36	
Asp[3]	8.17	4.57	2.645 2.618			
Trp[4]	7.81	4.40	3.04			$H_{\delta1}$ 7.00 $H_{\varepsilon1}$ 9.67 $H_{\varepsilon3}$ 7.29 $H_{\zeta3}$ 6.76 $H_{\zeta2}$ 9.69 H_{η} 6.81
Asp[5]	7.37	4.48	2.29			
Pro[6]		3.96	2.13	1.85	3.61 3.43	
Thr[7]	7.57	3.72	3.85	0.98		
His[8]	7.12	2.31	1.25 1.08			$H_{\delta1}$ 9.11 $H_{\delta2}$ -0.382 $H_{\varepsilon1}$ 0.024
Gln[9]	7.23	3.25	1.76	2.10		$H_{\varepsilon12}$ 7.31 6.28
Glu[10]	7.38	3.87	1.97 1.92	2.30		
Lys[11]	7.15	3.79	0.14	1.17	1.42	$H_{\varepsilon,\varepsilon'}$ 2.78 2.18 H_{ζ} 7.88
Leu[12]	7.22	4.28	1.50 1.07	1.50	1.78 1.30	
Glu[13]	8.16	3.89	2.1	2.50 2.27		
Lys[14]	8.10	4.06	2.01	1.53	1.73	$H_{\varepsilon,\varepsilon'}$ 3.02
Arg[15]	8.00	4.39	2.48 2.36	2.01	3.58 3.46	H_{ε} 7.55
Leu[16]	8.47	4.39	2.00	1.66	1.06 1.03	
Lys[17]	8.09	4.31	1.97	1.52	1.65	$H_{\varepsilon,\varepsilon'}$ 3.00

Peptide chain linked to propionate at position 18						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
Thr[1]	7.91	4.37	4.27	1.20		
Asn[2]	8.35	4.69	2.65		6.62 7.36	
Asp[3]	8.17	4.57	2.64 2.62	8.17	4.57	
Trp[4]	7.81	4.40	3.04			$H_{\delta 1}$ 7.00 $H_{\varepsilon 1}$ 6.69 $H_{\varepsilon 3}$ 7.29 $H_{\zeta 3}$ 6.76 $H_{\zeta 2}$ 9.96 H_{η} 6.85
Asp[5]	7.34	4.51	2.27			
Pro[6]		3.98	2.13	1.85	3.61 3.43	
Thr[7]	7.55	3.72	3.83	0.97		
His[8]	7.08	2.37	1.27 1.09			$H_{\delta 1}$ 9.17 $H_{\delta 2}$ -0.40 $H_{\varepsilon 1}$ 0.092
Gln[9]	7.28	3.31	1.78	2.12		$H\varepsilon_{12}$ 7.32 6.31
Glu[10]	7.46	3.88	1.98 1.91	2.30 2.22		
Lys[11]	7.20	3.79	0.16	1.14	1.42	$H\varepsilon_{,\varepsilon'}$ 2.68 2.18 H_{ζ} 7.77
Leu[12]	7.29	4.17	1.48 1.24	1.38	0.75 0.55	
Glu[13]	8.10	3.90	2.09	2.49 2.27		
Lys[14]	8.05	4.06	2.00	1.53	1.71	Hε,ε' 3.00
Arg[15]	7.92	4.37	2.36 2.29	2.01	3.58 3.46	Hε 7.55
Leu[16]	8.39	4.36	1.95	1.64	1.03 1.02	
Lys[17]	8.04	4.31	1.96	1.52	1.65	

Deuteroporphyrin IX assignments.			
2 -α,α'CH_2	4.86, 4.26	8H	9.42
18-α,α'CH_2	4.87, 4.27	10 H	10.48
2 -β,β'CH_2	3.73, 3.36	12CH3	3.76
18-β,β'CH_2	3.68, 3.33	13 H	9.35
3CH3	3.73	15 H	10.34
5H	10.5	17CH3	3.68
7 CH3	3.78	20 H	10.41

Table 13: Co(III)-Mimochrome VII **compound 2** assignment.						
Peptide chain linked to propionate at position 2						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
Thr[1]						
Asn[2]	8.38	4.74	2.69		6.66 7.43	
Asp[3]	8.23	4.63	2.72			
Trp[4]	7.79	4.3	2.93 2.88			$H_{\delta1}$ 6.93 $H_{\varepsilon1}$ 9.81 $H_{\varepsilon3}$ 7.08 $H_{\zeta3}$ 7.08 $H_{\zeta2}$ 7.21 H_{η} 6.49
Asp[5]	7.03	4.5	2.39			
Pro[6]		3.91	2.17	1.74	3.56 3.47	
Thr[7]	7.45	3.61	3.55	0.78		
His[8]	6.88	2.26	1.08 0.71			$H_{\delta1}$ 9.44 $H_{\delta2}$ -0.507 $H_{\varepsilon1}$ 0.155
Gln[9]	7.33	3.54	1.87	2.19		Hε,ε' 7.40 6.40
Glu[10]	7.65	3.91	1.95 1.91	2.32		
Lys[11]	7.27	3.81	1.05	-0.15	1.04	$H_{\varepsilon,\varepsilon'}$ 3.13 2.06 Hζ 6.83
Leu[12]	7.61	4.23	1.60	1.74 1.56	1.02 0.83	
Glu[13]	8.18	3.91	2.11	2.54 2.30		
Lys[14]	8.00	4.06	1.95	1.51	1.72	
Arg[15]	7.94	4.31	2.28 2.24	1.96	3.49 3.39	Hε 7.49
Leu[16]	8.39					
Lys[17]						

Peptide chain linked to propionate at position 18						
Aa	**H-N**	**H-α**	**H-β**	**H-γ**	**H-δ**	**others**
Thr[1]						
Asn[2]	8.35	4.69	2.65		6.65 7.40	
Asp[3]	8.19	4.7	2.65			
Trp[4]	7.75	4.3	2.90 2.81			$H_{\delta 1}$ 6.94 $H_{\varepsilon 1}$ 9.73 $H_{\varepsilon 3}$ 7.17 $H_{\zeta 3}$ 7.14 $H_{\zeta 2}$ 7.28 H_{η} 6.51
Asp[5]	7.18	4.48	2.4			
Pro[6]		3.91	2.17	1.74	3.56 3.47	
Thr[7]	7.39	3.59	3.44	0.73		
His[8]	6.84	2.34	1.16 0.82			$H_{\delta 1}$ 9.65 $H_{\delta 2}$ -0.504 $H_{\varepsilon 1}$ 0.20
Gln[9]	7.43	3.49	1.86	2.16		$H_{\varepsilon 21}$ 7.33 $H_{\varepsilon 22}$ 6.41
Glu[10]	7.75	3.90	1.95 1.87	2.34		
Lys[11]	7.28	3.79	1.09	0.14	1.10	$H_{\varepsilon,\varepsilon'}$ 3.17 2.14 H_{ζ} 9.96
Leu[12]	7.64	4.13	1.59 1.56	1.44	0,93 0.56	
Glu[13]	8.16	3.89	2.13			
Lys[14]	7.98	4.03	1.97	1.52	1.70	
Arg[15]	7.86	4.28	2.20 2.16	1.92	3.45 3.35	Hε 7.46
Leu[16]	8.33	4.31	1.90	1.61	0.97	
Lys[17]	7.99	4.28	1.93	1.50	1.64	

Deuteroporphyrin IX assignments.			
2 -α,α'CH_2	4.79, 4.37	8H	9.45
18-α,α'CH_2	4.77, 4.33	10 H	10.59
2 -β,β'CH_2	3.33, 3.26	12CH3	3.82
18-β,β'CH_2	3.33, 3.26	13 H	9.31
3CH3	3.68	15 H	10.34
5H	10.50	17CH3	3.62
7 CH3	3.74	20 H	10.43

Intensities of dipolar connectivities in the two-dimensional NOESY spectrum, obtained using a 200 ms mixing time, were measured using the integration subroutine of the CARA program. Peak volumes were converted into upper limit of inter-atomic distances, to be used as input for structural calculation.

A total of 777 NOESY cross peaks was assigned, integrated and transformed in upper distance limits with the program CALIBA,[158] for peak a and peak b, respectively. Of these NOE upper distance limits, 177 were removed, because they do not effectively restrict the conformation. In conclusion, we obtain two set of 554 upper distance limit restraints that were used as input for structural calculation. As for Co(III)-Mimochrome VI-2U1L in order to perform the simulated annealing, two non standard residue were added to the CYANA standard library and additional restraints were added to simulate the histidine coordination, and the amide bond formation between the lysine side chain and the deuterporphyrin. The quality of the obtained structures was evaluated in terms of deviations from ideal lengths and bond angles through a Ramachandran plot obtained using the program PROCHECK-NMR.

8. Fe(III)-Mimochrome IV

8.1. NMR data collection

The NMR analysis of the paramagnetic Fe(III)-Mimochrome IV complex was performed in phosphate buffer (10 mM, pH 6.5)/TFE/DMSO solution (60/20/20, v/v/v) at 298 K. NMR samples were prepared by dissolving weighted amounts of the compound in the solvent system (V = 0.600 ml) for a final concentration of 0.6 mM. Proton 1D and 2D nuclear overhauser spectra were collected on a Bruker Avance II 500 MHz spectrometer with a 5 mm QXI probe at "Centro de Ressonância Magnética António Xavier" (CERMAX), hosted at the Instituto de Tecnologia Química e Biológica (ITQB), Oeiras, Portugal.

Proton ^{1}H nmr spectra were acquired with a spectral width of 30 kHz, collecting 4 K data points. The water signal was suppressed by using a 500 ms selective pulse.

WEFT-NOESY spectrum was performed as described in the literature with an acquisition time of 50 ms, an interpulse delay (t) of 50 ms and a mixing time of 50 ms. The spectral width was of 30 kHz in both dimension. The data file generally consisted of 4K and 512 data points in F2 and F1, respectively.

8.2. Determination of the axial ligand orientations

The axial ligand orientations of Fe(III)-Mimochrome IV was determined by using the empirical equation proposed by Turner:

$$\delta_i(\text{ppm}) = \cos\beta[38.0 \sin^2(\theta_i-\phi)-4.1 \cos^2(\theta_i+\phi)-15.9]+13.8 \qquad \textbf{(eq.7)}$$

This equation contains two unknown variables, independent from each other,

that define the axial ligand orientations:

- β, the acute angle between the two histidine planes, and ϕ, the average orientation of the His planes, projected on the heme plane, with respect to the N_{21}-N_{23} direction (see figure 3 in the main text).
- δ_i is the hyperfine shift of the *i*th methyl, experimentally determine for each methyl from NMR spectra.
- θ_i, is the angle between the metal-*i*th methyl direction and the metal-N_{23} axis; it represents a parameter which is proper to each methyl, and can be deduced from the heme geometry. In this paper, we used the θ_i values proposed by Turner (table 14).

The equation (7) was implemented in Microsoft office excel and, for each methyl, the experimental δ value and the θ value, as derived from Turner (see table 13), were inserted. The β and ϕ values were optimized to obtain the best fit to the four experimental δ_i values (see table 15).

Table 14 θ_i for the heme methyl groups

methyl	θ_i
7-CH_3	108
12-CH_3	18
17-CH_3	-80
3-CH_3	162

Table 15 ^{1}H chemical shifts of the heme methyl groups

methyl	δ (ppm) experimental	δ (ppm) calculated
7-CH_3	26.1	23.4
12-CH_3	6.7	4.1
17-CH_3	24.5	24.8
3-CH_3	16.9	17.0

List of acronyms and abbreviations

Abs	Absorbance
ABTS	2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid)
Ac	Acetyl
Aib	α-aminoisobutyric acid
Ala	Alanine
a.m.u.	Atomic mass unit
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
BOC	t-Butoxycarbonyl
CD	Circular Diichroism
CCP	Cytochrome *c* peroxidase
Cys	Cysteine
cyt	Cytochrome
DCM	Dichloromethane
DIEA	Diisopropylethylamine
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DP IX	Deuteroporphyrin IX
DQF-COSY	Double Quantum Filter- COrrelation SpectroscopY
EDT	1,2-ethanedithiol
ESI -MS	Electrospray Ionization- Mass Spectrometry
far-UV	Far-Ultraviolet
Fmoc	9-Fluorenylmethoxycarbonyl
Gln	Glutamine

Glu	Glutamic acid
Gly	Glycine
HATU	N-[(dimethylamino)-1H-1,2,3-triazolo [4,5-b]pyridin-1-yl-methylene]-N methylmethanaminium hexafluorophosphate N-oxide
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium
His	Histidine
HOBt	N-hydroxybenzotriazole
HRP	Horseradish Peroxidase
HSAB	Hard and soft acids and bases theory
ICP-MS	Inductively Coupled Plasma - Mass Spectrometry
Ile	Isoleucine
IT-TOF	Ion Trap- Time Of Flight
kcat	Turnover number (in the Michaelis-Menten equation)
LC	Liquid Chromatography
Leu	Leucine
Lys	Lysine
Met	Methionine
Mmt	Methoxytrityl
MP 8	Microperoxidases 8
MP 11	Microperoxidases 11
MS	Mass Spectrometry
NMP	N-Methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser SpectroscopY
Np	Heme pyrrole nitrogen
OtBu	Tert-Butoxy

Pbf	2,2,4,6,7-Pentamethyldihydrobenzofurane-5-sulfonyl
Pdb	Protein Data Bank
ppm	Part per million
Pro	Proline
RP-HPLC	Reverse Phase - High Pressure Liquid Chromatography
Ser	Serine
SPPS	Solid Phase Peptide Synthesis
Rmsd	Root-mean-square deviation
tBu	Tert-Butyl
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
Thr	Threonine
TLC	Thin Layer Chromatography
TOCSY	TOtal Correlated SpectroscopY
Trp	Tryptophan
Trt	Trityl
Tyr	Tyrosine
UV-vis	Ultra Violet/Visible
VMD	Visual Molecular Dynamics
Xaa	Unspecified Amino Acid

References

1. Y. Lu, N. Yeung, N. Sieracki, and N. M. Marshall, *Nature*, 2009, **460**, 855–862.
2. K. J. Waldron, J. C. Rutherford, D. Ford, and N. J. Robinson, *Nature*, 2009, **460**, 823–830.
3. O. Maglio, F. Nastri, and A. Lombardi, in *Ionic Interactions in Natural and Synthetic Macromolecules*, eds. A. Ciferri and A. Perico, John Wiley & Sons, Inc., 2012, pp. 361–450.
4. S. K. Chapman, S. Daff, and A. W. Munro, in *Metal Sites in Proteins and Models*, eds. H. A. O. Hill, P. J. Sadler, and A. J. Thomson, Springer Berlin Heidelberg, 1997, pp. 39–70.
5. C. J. Reedy and B. R. Gibney, *Chem. Rev.*, 2004, **104**, 617–650.
6. A. Lombardi, F. Nastri, and V. Pavone, *Chem. Rev.*, 2001, **101**, 3165–3189.
7. F. Nastri, L. Lista, P. Ringhieri, R. Vitale, M. Faiella, C. Andreozzi, P. Travascio, O. Maglio, A. Lombardi, and V. Pavone, *Chem. – Eur. J.*, 2011, **17**, 4444–4453.
8. C. Andreini, I. Bertini, G. Cavallaro, G. L. Holliday, and J. M. Thornton, *J. Biol. Inorg. Chem.*, 2008, **13**, 1205–1218.
9. H. B. Gray and J. R. Winkler, *Annu. Rev. Biochem.*, 1996, **65**, 537–561.
10. S. J. Lippard and J. M. Berg, *Principles of bioinorganic chemistry*, University Science Books, Mill Valley, Calif., 1994.
11. I. Bertini, *Biological Inorganic Chemistry: Structure and Reactivity*, University Science Books, 2007.
12. C. W. Bock, A. K. Katz, G. D. Markham, and J. P. Glusker, *J. Am. Chem. Soc.*, 1999, **121**, 7360–7372.
13. R. R. Crichton and J. L. Pierre, *Biometals*, 2001, **14**, 99–112.
14. R. J. P. Williams, *Chem. Commun.*, 2003, 1109–1113.
15. T. Dudev and C. Lim, *Chem. Rev.*, 2003, **103**, 773–788.
16. J. A. Cowan, *Chem. Rev.*, 1998, **98**, 1067–1088.
17. Y. Zhang and V. N. Gladyshev, *J. Biol. Chem.*, 2010, **285**, 3393–3405.
18. R. J. P. Williams, *Cell. Mol. Life Sci.*, 1997, **53**, 816–829.
19. R. H. Holm, P. Kennepohl, and E. I. Solomon, *Chem. Rev.*, 1996, **96**, 2239–2314.
20. S. M. Yannone, S. Hartung, A. L. Menon, W. W. Adams, and J. A. Tainer, *Curr. Opin. Biotechnol.*, 2012, **23**, 89–95.
21. I. A. Kaltashov, M. Zhang, S. J. Eyles, and R. R. Abzalimov, *Anal. Bioanal. Chem.*, 2006, **386**, 472–481.
22. A. S. Borovik, *Acc. Chem. Res.*, 2005, **38**, 54–61.
23. S. M. Berry, M. H. Baker, and N. J. Reardon, *J. Inorg. Biochem.*, 2010, **104**, 1071–1078.
24. R. Yerushalmi, A. Brandis, V. Rosenbach-Belkin, K. K. Baldridge, and A. Scherz, *J. Phys. Chem. A*, 2006, **110**, 412–421.
25. R. L. Shook and A. S. Borovik, *Inorg. Chem.*, 2010, **49**, 3646–3660.

26. R. G. Pearson, *J. Am. Chem. Soc.*, 1963, **85**, 3533–3539.
27. K. Degtyarenko, *Bioinformatics*, 2000, **16**, 851–864.
28. K. Degtyarenko and S. Contrino, *BMC Struct. Biol.*, 2004, **4**, 3.
29. I. D. Campbell, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 377–381.
30. H. M. Berman, *Acta Crystallogr. A*, 2008, **64**, 88–95.
31. F. He, C. L. Hendrickson, and A. G. Marshall, *J. Am. Soc. Mass Spectrom.*, 2000, **11**, 120–126.
32. K. A. Johnson, M. F. J. M. Verhagen, P. S. Brereton, M. W. W. Adams, and I. J. Amster, *Anal. Chem.*, 2000, **72**, 1410–1418.
33. K. A. Johnson, B. A. Shira, J. L. Anderson, and I. J. Amster, *Anal. Chem.*, 2001, **73**, 803–808.
34. C. Andreini, I. Bertini, and A. Rosato, *Accounts Chem. Res.*, 2009, **42**, 1471–1479.
35. D. E. Fenton, *Biocoordination Chemistry*, Oxford University Press, Oxford, 1995.
36. M. Vidakovic, S. G. Sligar, H. Li, and T. L. Poulos, *Biochemistry*, 1998, **37**, 9211–9219.
37. L. Banci, I. Bertini, H. B. Gray, C. Luchinat, T. Reddig, A. Rosato, and P. Turano, *Biochemistry*, 1997, **36**, 9867–9877.
38. M. Gajhede, D. J. Schuller, A. Henriksen, A. T. Smith, and T. L. Poulos, *Nat. Struct. Mol. Biol.*, 1997, **4**, 1032–1038.
39. R. Liddington, Z. Derewenda, E. Dodson, R. Hubbard, and G. Dodson, *J. Mol. Biol.*, 1992, **228**, 551–579.
40. J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips, *Nature*, 1958, **181**, 662–666.
41. M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, and A. C. T. North, *Nature*, 1960, **185**, 416–422.
42. J.C. Kendrew, *Nobel Lecture, December 11, 1962*.
43. L. J. Smith, A. Kahraman, and J. M. Thornton, *Proteins*, 2010, **78**, 2349–2368.
44. B. K. Vainshtein, W. R. Melik-Adamyan, V. V. Barynin, A. A. Vagin, A. I. Grebenko, V. V. Borisov, K. S. Bartels, I. Fita, and M. G. Rossmann, *J. Mol. Biol.*, 1986, **188**, 49–61.
45. Buchler, *The Porphyrins*, Academic Press, New York, 1979, vol. 1.
46. T. M. Iverson, D. M. Arciero, B. T. Hsu, M. S. P. Logan, A. B. Hooper, and D. C. Rees, *Nat. Struct. Biol.*, 1998, **5**, 1005–1012.
47. R. E. Dickerson, T. Takano, D. Eisenberg, O. B. Kallai, L. Samson, A. Cooper, and E. Margoliash, *J. Biol. Chem.*, 1971, **246**, 1511–1535.
48. S. C. Andrews, P. M. Harrison, S. J. Yewdall, P. Arosio, S. Levi, W. Bottke, M. von Darl, J.-F. Briat, J.-P. Laulhère, and S. Lobreaux, *J. Inorg. Biochem.*, 1992, **47**, 161–174.
49. T. L. Poulos, B. C. Finzel, and A. J. Howard, *J. Mol. Biol.*, 1987, **195**, 687–700.
50. P. Kotsonis, L. G. Fröhlich, C. S. Raman, H. Li, M. Berg, R. Gerwig, V. Groehn, Y. Kang, N. Al-Masoudi, S. Taghavi-Moghadam, D. Mohr, U.

Münch, J. Schnabel, P. Martásek, B. S. S. Masters, H. Strobel, T. Poulos, H. Matter, W. Pfleiderer, and H. H. H. W. Schmidt, *J. Biol. Chem.*, 2001, **276**, 49133–49141.
51. O. Einsle, A. Messerschmidt, P. Stach, G. P. Bourenkov, H. D. Bartunik, R. Huber, and P. M. H. Kroneck, *Nature*, 1999, **400**, 476–480.
52. D. Leys, K. Backers, T. E. Meyer, W. R. Hagen, M. A. Cusanovich, and J. J. V. Beeumen, *J. Biol. Chem.*, 2000, **275**, 16050–16056.
53. P. Arnoux, R. Haser, N. Izadi, A. Lecroisey, M. Delepierre, C. Wandersman, and M. Czjzek, *Nat. Struct. Mol. Biol.*, 1999, **6**, 516–520.
54. M. F. Perutz, P. D. Pulsinelli, and H. M. Ranney, *Nature*, 1972, **237**, 259–263.
55. S. E. Martinez, D. Huang, M. Ponomarev, W. A. Cramer, and J. L. Smith, *Protein Sci.*, 1996, **5**, 1081–1092.
56. W. N. Lanzilotta, D. J. Schuller, M. V. Thorsteinsson, R. L. Kerby, G. P. Roberts, and T. L. Poulos, *Nat. Struct. Mol. Biol.*, 2000, **7**, 876–880.
57. F. A. Walker, . Boi Hanh Huynh, W. R. Scheidt, and S. R. Osvath, *J. Am. Chem. Soc.*, 1986, **108**, 5288–5297.
58. A. M. Azevedo, V. C. Martins, D. M. F. Prazeres, V. Vojinović, J. M. S. Cabral, and L. P. Fonseca, in *Biotechnology Annual Review*, Elsevier, 2003, vol. Volume 9, 199–247.
59. N. Fawal, Q. Li, B. Savelli, M. Brette, G. Passaia, M. Fabre, C. Mathé, and C. Dunand, *Nucleic Acids Res.*, 2013, **41**, D441–D444.
60. T. L. Poulos, S. T. Freer, R. A. Alden, S. L. Edwards, U. Skogland, K. Takio, B. Eriksson, N. Xuong, T. Yonetani, and J. Kraut, *J. Biol. Chem.*, 1980, **255**, 575–580.
61. T. L. Poulos, *Arch. Biochem. Biophys.*, 2010, **500**, 3–12.
62. C. A. Bonagura, B. Bhaskar, H. Shimizu, H. Li, M. Sundaramoorthy, D. E. McRee, D. B. Goodin, and T. L. Poulos, *Biochemistry* , 2003, **42**, 5600–5608.
63. S. P. Cramer, J. H. Dawson, K. O. Hodgson, and L. P. Hager, *J. Am. Chem. Soc.*, 1978, **100**, 7282–7290.
64. J. N. Rodríguez-López, D. J. Lowe, J. Hernández-Ruiz, A. N. P. Hiner, F. García-Cánovas, and R. N. F. Thorneley, *J. Am. Chem. Soc.*, 2001, **123**, 11838–11847.
65. L. Banci, *J. Biotechnol.*, 1997, **53**, 253–263.
66. M. I. Savenkova, S. L. Newmyer, and P. R. O. de Montellano, *J. Biol. Chem.*, 1996, **271**, 24598–24603.
67. N. C. Veitch, *Phytochemistry*, 2004, **65**, 249–259.
68. L. B. Vitello, J. E. Erman, M. A. Miller, J. Wang, and J. Kraut, *Biochemistry*, 1993, **32**, 9807–9818.
69. C. A. Mac Munn, *Philos. Trans. R. Soc. Lond.*, 1886, **177**, 267–298.
70. D. Keilin, *Proc. R. Soc. Lond. B*, 1925, **98**, 312-339.
71. B. Chance, A. Weber, *J. Physiol*. 1963, **169**, 263-277. .
72. B. A. Haddock and C. W. Jones, *Bacteriol. Rev.*, 1977, **41**, 47–99.
73. P. F. Urban and M. Klingenberg, *Eur. J. Biochem.*, 1969, **9**, 519–525.

74. O. Vallon, L. Bulte, P. Dainese, J. Olive, R. Bassi, and F. A. Wollman, *Proc. Natl. Acad. Sci.*, 1991, **88**, 8262–8266.
75. G. Vergères and L. Waskell, *Biochimie*, 1995, **77**, 604–620.
76. A. K. Churg and A. Warshel, *Biochemistry*, 1986, **25**, 1675–1681.
77. M. L. Zastrow and V. L. Pecoraro, *Coord. Chem. Rev.*, 2013, **257**, 2565–2588.
78. S. J. Lippard, *Nat. Chem. Biol.*, 2006, **2**, 504–507.
79. G. Xing and V. J. DeRose, *Curr. Opin. Chem. Biol.*, 2001, **5**, 196–200.
80. D. Ghosh and V. L. Pecoraro, *Curr. Opin. Chem. Biol.*, 2005, **9**, 97–103.
81. M. A. Case and G. L. McLendon, *Acc. Chem. Res.*, 2004, **37**, 754–762.
82. Y. Lu, *Inorg. Chem.*, 2006, **45**, 9930–9940.
83. M. L. Kennedy and B. R. Gibney, *Curr. Opin. Struct. Biol.*, 2001, **11**, 485–490.
84. M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, and A. Lombardi, *Nat. Chem. Biol.*, 2009, **5**, 882–884.
85. R. T. M. de Rosales, M. Faiella, E. Farquhar, L. Que Jr, C. Andreozzi, V. Pavone, O. Maglio, F. Nastri, and A. Lombardi, *J. Biol. Inorg. Chem.*, 2010, **15**, 717–728.
86. H. W. Hellinga, *Curr. Opin. Biotechnol.*, 1996, **7**, 437–441.
87. Y. Lu, S. M. Berry, and T. D. Pfister, *Chem. Rev.*, 2001, **101**, 3047–3080.
88. M. L. Kennedy, S. Silchenko, N. Houndonougbo, B. R. Gibney, P. L. Dutton, K. R. Rodgers, and D. R. Benson, *J. Am. Chem. Soc.*, 2001, **123**, 4635–4636.
89. F. Nastri, R. Bruni, O. Maglio, and A. Lombardi, in *Coordination Chemistry in Protein Cages*, eds. T. Ueno and Y. Watanabe, John Wiley & Sons, Inc., 2013, pp. 43–85.
90. C. D. Coldren, H. W. Hellinga, and J. P. Caradonna, *Proc. Natl. Acad. Sci.*, 1997, **94**, 6635–6640.
91. N. Yeung, Y.-W. Lin, Y.-G. Gao, X. Zhao, B. S. Russell, L. Lei, K. D. Miner, H. Robinson, and Y. Lu, *Nature*, 2009, **462**, 1079–1082.
92. A. Lombardi, D. Marasco, O. Maglio, L. Di Costanzo, F. Nastri, and V. Pavone, *Proc. Natl. Acad. Sci.*, 2000, **97**, 11922–11927.
93. W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, and A. Lombardi, *Annu. Rev. Biochem.*, 1999, **68**, 779–819.
94. S. Chakraborty, J. Yudenfreund Kravitz, P. W. Thulstrup, L. Hemmingsen, W. F. DeGrado, and V. L. Pecoraro, *Angew. Chem. Int. Ed.*, 2011, **50**, 2049–2053.
95. P. D. Barker, *Curr. Opin. Struct. Biol.*, 2003, **13**, 490–499.
96. A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, and W. F. DeGrado, *Proc. Natl. Acad. Sci.*, 2000, **97**, 6298–6305.
97. L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, and A. Lombardi, *J. Am. Chem. Soc.*, 2001, **123**, 12749–12757.
98. J. R. Calhoun, F. Nastri, O. Maglio, V. Pavone, A. Lombardi, and W. F. DeGrado, *Biopolymers*., 2005, **80**, 264–278.

99. O. Maglio, F. Nastri, R. T. Martin de Rosales, M. Faiella, V. Pavone, W. F. DeGrado, and A. Lombardi, *Comptes Rendus Chim.*, 2007, **10**, 703–720.
100. D. R. Benson, B. R. Hart, X. Zhu, and M. B. Doughty, *J. Am. Chem. Soc.*, 1995, **117**, 8502–8510.
101. P. A. Arnold, D. R. Benson, D. J. Brink, M. P. Hendrich, G. S. Jas, M. L. Kennedy, D. T. Petasis, and M. Wang, *Inorg. Chem.*, 1997, **36**, 5306–5315.
102. A. B. Cowley, M. L. Kennedy, S. Silchenko, G. S. Lukat-Rodgers, K. R. Rodgers, and D. R. Benson, *Inorg. Chem.*, 2006, **45**, 9985–10001.
103. S. Shinde, J. M. Cordova, B. W. Woodrum, and G. Ghirlanda, *J. Biol. Inorg. Chem.*, 2012, **17**, 557–564.
104. M. Faiella, O. Maglio, F. Nastri, A. Lombardi, L. Lista, W. R. Hagen, and V. Pavone, *Chem. – Eur. J.*, 2012, **18**, 15960–15971.
105. R. L. Koder, J. L. R. Anderson, L. A. Solomon, K. S. Reddy, C. C. Moser, and P. L. Dutton, *Nature*, 2009, **458**, 305–309.
106. H. M. Marques, *Inorg. Chem.*, 2005, **44**, 6146–6148.
107. D. Verbaro, A. Hagarman, A. Kohli, and R. Schweitzer-Stenner, *J. Biol. Inorg. Chem.*, 2009, **14**, 1289–1300.
108. S. Mathura, D. Sannasy, A. S. de Sousa, C. B. Perry, I. Navizet, and H. M. Marques, *J. Inorg. Biochem.*, 2013, **123**, 66–79.
109. S. Nicolis, L. Casella, R. Roncone, C. Dallacosta, and E. Monzani, *Comptes Rendus Chim.*, 2007, **10**, 380–391.
110. O. Q. Munro and H. M. Marques, *Inorg. Chem.*, 1996, **35**, 3752–3767.
111. H. M. Marques, *Dalton Trans.*, 2007, 4371–4385.
112. M. H. Hecht, K. M. Vogel, T. G. Spiro, N. R. L. Rojas, S. Kamtekar, C. T. Simons, J. E. Mclean, and R. S. Farid, *Protein Sci.*, 1997, **6**, 2512–2524.
113. Y. Wei, S. Kim, D. Fela, J. Baum, and M. H. Hecht, *Proc. Natl. Acad. Sci.*, 2003, **100**, 13270–13273.
114. A. Go, S. Kim, J. Baum, and M. H. Hecht, *Protein Sci.*, 2008, **17**, 821–832.
115. S. C. Patel and M. H. Hecht, *Protein Eng. Des. Sel.*, 2012, **25**, 445–451.
116. L. Casella, M. Gullotti, E. Monzani, L. De Giola, and F. Chillemi, *Rend. Fis. Acc. Lincei*, 1991, **2**, 201–212.
117. M. Faiella, O. Maglio, F. Nastri, A. Lombardi, L. Lista, W. R. Hagen, and V. Pavone, *Chem. - Eur. J.*, 2012, **18**, 15960–15971.
118. F. Nastri, A. Lombardi, G. Morelli, O. Maglio, G. D'Auria, C. Pedone, and V. Pavone, *Chem. – Eur. J.*, 1997, **3**, 340–349.
119. G. D'Auria, O. Maglio, F. Nastri, A. Lombardi, M. Mazzeo, G. Morelli, L. Paolillo, C. Pedone, and V. Pavone, *Chem. – Eur. J.*, 1997, **3**, 350–362.
120. F. Nastri, A. Lombardi, G. Morelli, C. Pedone, V. Pavone, G. Chottard, P. Battioni, and D. Mansuy, *J. Biol. Inorg. Chem.*, 1998, **3**, 671–681.
121. C. N. Pace and J. M. Scholtz, *Biophys. J.*, 1998, **75**, 422–427.
122. A. Lombardi, F. Nastri, M. Sanseverino, O. Maglio, C. Pedone, and V. Pavone, *Inorganica Chim. Acta*, 1998, **275–276**, 301–313.
123. A. Lombardi, F. Nastri, D. Marasco, O. Maglio, G. De Sanctis, F. Sinibaldi, R. Santucci, M. Coletta, and V. Pavone, *Chem. – Eur. J.*, 2003, **9**, 5643–5654.

124. L. Di Costanzo, S. Geremia, L. Randaccio, F. Nastri, O. Maglio, A. Lombardi, and V. Pavone, *J. Biol. Inorg. Chem.*, 2004, **9**, 1017–1027.
125. J. N. Rodriguez-Lopez, A. T. Smith, and R. N. F. Thorneley, *J. Biol. Chem.*, 1996, **271**, 4023–4030.
126. I. L. Karle and P. Balaram, *Biochemistry*, 1990, **29**, 6747–6756.
127. M. Bellanda, E. Peggion, S. Mammi, R. Bürgi, and W. Van Gunsteren, *J. Pept. Res.*, 2001, **57**, 97–106.
128. J. Jeneer, B.H. Meier, P. Bachmann, R.R.Ernst, *J. Chem. Phys*., 1979, **71**, 4546-4553.
129. A. Bax and D. G. Davis, *J. Magn. Reson.*, 1985, **65**, 355–360.
130. U. Piantini, O. W. Sorensen, and R. R. Ernst, *J. Am. Chem. Soc.*, 1982, **104**, 6800–6801.
131. K. Wuthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.
132. D. S. Wishart, B. D. Sykes, and F. M. Richards, *J. Mol. Biol.*, 1991, **222**, 311–333.
133. R. D. Guiles, J. Altman, I. D. Kuntz, and L. Waskell, *Biochemistry*, 1990, **29**, 1276–1289.
134. *A. Pastore, V. Saudek. J. Magn. Reson.,* 1990, **90**, 165-176
135. P. Güntert, in *Protein NMR Techniques*, ed. A. K. Downing, Humana Press, 2004, 353–378.
136. R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
137. S. S. Huang, R. L. Koder, M. Lewis, A. J. Wand, and P. L. Dutton, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 5536–5541.
138. S. Vuilleumier and M. Mutter, *Biopolymers*, 1993, **33**, 389–400.
139. E. K. S. Vijayakumar, T. S. Sudha, and P. Balaram, *Biopolymers*, 1984, **23**, 877–886.
140. A. W. Johnson and I. T. Kay, *J. Chem. Soc. Resumed*, 1960, 2979–2983.
141. M. Momenteau and D. Lexa, *Biochem. Biophys. Res. Commun.*, 1974, **58**, 940–944.
142. O. R. Fennema, *Food Chemistry, Third Edition*, CRC Press, 1996.
143. G. Otting, *Annu. Rev. Biophys.*, 2010, **39**, 387–405.
144. I. Bertini, P. Turano, and A. J. Vila, *Chem. Rev.*, 1993, **93**, 2833–2932.
145. D. L. Turner, *J. Biol. Inorg. Chem.*, 2000, **5**, 328–332.
146. I. Bertini, C. Luchinat, G. Parigi, and F. A. Walker, *J. Biol. Inorg. Chem.*, 1999, **4**, 515–519.
147. N. V. Shokhirev and F. A. Walker, *J. Biol. Inorg. Chem.*, 1998, **3**, 581–594.
148. R. O. Louro, I. J. Correia, L. Brennan, I. B. Coutinho, A. V. Xavier, and D. L. Turner, *J. Am. Chem. Soc.*, 1998, **120**, 13240–13247.
149. J. M. Dantas, I. H. Saraiva, L. Morgado, M. A. Silva, M. Schiffer, C. A. Salgueiro, and R. O. Louro, *Dalton Trans.*, 2011, **40**, 12713–12718.
150. A. P. Aguiar, H. S. Costa, R. O. Louro, A. V. Xavier, and D. L. Turner, *Inorganica Chim. Acta*, 1998, **273**, 196–200.

151. J. M. Dantas, L. Morgado, Y. Y. Londer, A. P. Fernandes, R. O. Louro, P. R. Pokkuluri, M. Schiffer, and C. A. Salgueiro, *J. Biol. Inorg. Chem.*, 2012, **17**, 11–24.
152. B. M. Fonseca, I. H. Saraiva, C. M. Paquete, C. M. Soares, I. Pacheco, C. A. Salgueiro, and R. O. Louro, *J. Biol. Inorg. Chem.*, 2009, **14**, 375–385.
153. L. Morgado, I. H. Saraiva, R. O. Louro, and C. A. Salgueiro, *FEBS Lett.*, 2010, **584**, 3442–3445.
154. D. W. Low, H. B. Gray, and J. Ø. Duus, *J. Am. Chem. Soc.*, 1997, **119**, 1–5.
155. F. Yang, T. K. Shokhireva, and F. A. Walker, *Inorg. Chem.*, 2011, **50**, 1176–1183.
156. W. Humphrey, A. Dalke, and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
157. Accelrys Software Inc., Discovery Studio Modeling Environment, Release 4.0, San Diego: Accelrys Software Inc., 2013.
158. P. Güntert, W. Braun, and K. Wüthrich, *J. Mol. Biol.*, 1991, **217**, 517–530.

Publications and communications

Claudia Vicari, Ivo H. Saraiva, Ornella Maglio, Flavia Nastri,Vincenzo Pavone, Ricardo O. Louro and Angela Lombardi.

Chem. Commun., 2014,**50**, 3852-3855.

Title:

"Artificial heme-proteins: determination of axial ligand orientations through paramagnetic NMR shifts"

Abstract:

An empirical equation, describing the relationship between the porphyrin methyl hyperfine shifts and the position of the axial ligand(s), has been applied to an artificial heme-protein in order to obtain insight into the active site properties of heme-protein models.

Claudia Vicari, Ornella Maglio, Concetta Andreozzi, Flavia Nastri, Vincenzo Pavone, Angela Lombardi.

XXVth International Conference on Magnetic Resonance in Biological Systems, Lyon, August 2012.

Title:

"Four-helix bundle motifs: the role of turns on folding and stability"

Claudia Vicari, Ornella Maglio, Concetta Andreozzi, Flavia Nastri, Vincenzo Pavone and Angela Lombardi.

Multidimensional NMR in Structural Biology EMBO pratical course, Joachimsthal, August 2012.

Title:

"Four-helix bundle motifs: the role of turns on folding and stability"

Claudia Vicari, Concetta Andreozzi, Daniela Liguori, Ornella Maglio, Flavia Nastri, Angela Lombardi, Vincenzo Pavone.

XXIV National Meeting of the Italian Chemical Society, Lecce, September 2011

Title:

" Design and Preliminary Characterization of New Heme-Protein Models"

Ornella Maglio, Claudia Vicari, Concetta Andreozzi, Flavia Nastri, Angela Lombardi,Vincenzo Pavone.

XXIV National Meeting of the Italian Chemical Society, Lecce, September 2011.

Title:

"Diiron-oxo protein models: the role of turns in stabilizing α- helical harpins"

Claudia Vicari, Concetta Andreozzi, Flavia Nastri, Vincenzo Pavone, Ornella Maglio, Angela Lombardi.

National school of Bioinorganic chemistry, Siena, July, 2011.

Title:

"Diiron-oxo protein models: the role of the loop in stabilizing four-helix bundle"

I wish to thank everyone whose support was essential to accomplish this project. In first place, all the people working in the Artificial Metallo Enzyme Group of Naples.

I'm deeply grateful to Prof. Vincenzo Pavone, who gave constant guidance at all stages of this research work. I am profoundly indebted to him for sharing his knowledge, for showing trust in my work, and for the patience he had in clarifying all my doubts and answering my innumerable questions.

I extend my thankfulness to Prof. Flavia Nastri, who encouraged me throughout this entire project, supporting in moment of difficulties, and helping to shape my ideas. Besides, My sincere gratitude go to her for her priceless help in writing this dissertation.

I wish to thanks to Prof. Angela Lombardi, for the support she always give to my studies, for her advice and encouragement. Her help was invaluable to me.

In a special way, I want to thanks Dott. Ornella Maglio. she co-operate with me in every step of this research project, with enormous patience and great interest. Without her help, this thesis would not be possible. There is no word I can explain my gratitude to her for continuous attention and unfailing concern.

I wish to thank Dr. Liliana Lista, for her wise advice and careful suggestions Also, I want to thanks her because during the time I spend writing this thesis, she never went home without ask how things were going.

I thanks Dr. Rosita Vitale, for being an important guide during these three years. She has often open my mind.

To Dr. Marco Chino, I'm in debt because he never deny help, no matter how much busy, or tired, he is.

To Dr. Corinne Cerrone, I want to express my thanks for all the time she help in all kinds of things, and for precious words each time I became crazy panicked.

To Dr. Daniela Liguori, Dr. Salvatore Costanzo, Dr. Giusy Genovese, Dr. Caterina Ferarra, Dr. Luisa Del Zoppo my acknowledgement are for their smiling, their suggestion and their kind word in the last, more difficult period of this thesis.

I also want to thanks some people that were working in the Artificial Metalloenzymes group in past time, especially Dr. Giorgio Caserta and Dr. Tina Andreozzi, whose help was essential to this work, and Dr. Rosa Bruni, Paolo Ringhieri, Manule Ruggiero, that share with me unforgettable moments.

I'm also deeply grateful to Prof. R.O: Louro, for its hospitality at the Inorganic Biochemistry and NMR group of the Instituto de Tecnologia Química e Biológica of Lisbona (Itqb). He give me the opportunity to increase my scientific background, to improve this work, and to spend a fantastic period in an exciting scientific environment.

Mi sincere thank goes to my examiner, Prof. Gerardino D'Errico. His interest in my work encouraged me to do better, and his advice increased my critical reasoning.

I also want to thanks my colleagues in this PhD course, for useful and amusing conversations we have, and Prof Luigi Paduano, for his concern in valorize our research.

To all this people, goes my gratitude for sharing with my their passion for science.

I want wish to thanks all my friend, because I always knew they will be still there, despite my disappearing.

To my family goes my biggest thanks. Their love and their support are the secret of my happiness.

Least buy not least, thanks to Andrea, because he loves me.